



A University of Sussex PhD thesis

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

*USING MASS SPECTROMETRY TO
QUANTIFY CHANGES IN HISTONE
POST-TRANSLATIONAL
MODIFICATIONS AS A RESULT OF DNA
DOUBLE-STRAND BREAKS*

Abubakar Abbas Hatimy

Thesis submitted for the degree of Doctor of Philosophy

University of Sussex

November 2018

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another university for the award of any other degree

Signature.....

Abubakar Abbas Hatimy

UNIVERSITY OF SUSSEX

ABUBAKAR ABBAS HATIMY

THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
BIOCHEMISTRY

USING MASS SPECTROMETRY TO QUANTIFY CHANGES IN HISTONE POST-
TRANSLATIONAL MODIFICATIONS AS A RESULT OF DNA DOUBLE-
STRANDED BREAKS

The structure of DNA is under constant threat from both extracellular and intracellular sources, causing mutations that could result in the reduced viability of the cell. In the nucleus of eukaryotes, DNA wraps around histone proteins 1.67 times. Two copies of Histone H2A, H2B, H3 and H4 congregate to form an octomer: along with the DNA strand together make the nucleosome. It has been shown that via histone post-translational modifications (HPTM), the nucleosome will signal to recruit biological machinery to process DNA such as during transcription or repair. One of the most characterised HPTM is phosphorylation of serine 139 of H2AX (γ H2AX), which is indicative of a DNA strand break — single or double. Subsequent to the phosphorylation, the cell's DNA damage response (DDR) will proceed to recruit proteins that repair the damage.

Recent studies have implicated a myriad of HPTMs in the DDR. However, many of the putative DDR HPTM were described with various immunoassays such as western blot and immunofluorescence. Immunoassays are notorious for reproducibility issues and experimental design biases. Comparatively, mass spectrometry is renowned for its precision and its ability to accommodate high throughput analysis of many datasets in an unbiased manner. We utilise MS to quantify γ H2AX changes after DNA damage and show that H2AX tyrosine 142 phosphorylation is not as abundant as previously shown. We also developed a method for simultaneously quantifying many HPTMs changes that occur as a result of the DDR at both proximal to the site of damage and a genomic scale. The developed method of quantifying DDR HPTM accommodates future experiments such as determining whether DDR HPTMs are old or new (with SILAC) or be adapted to analyse HPTMs as a result of transcription, translation or development. A translational aspect would be to quantify the epigenetic marks of cancer patients before and post-treatment.

Acknowledgments

Thomas Wayne: "And why do we fall, Bruce? So we can learn to pick ourselves up"

First and foremost, I would like to thank Steve for not just giving me the opportunity to undertake a PhD, but the endless patience and support he provided in both in and out of the lab. He always pushed you to do better because he knew you could. Steve's open-door policy was immensely helpful in receiving assistance and guidance to develop as a scientist by always asking new ways to scrutinise one's data. Again. Thank you.

I want to extend my gratitude to the other Sweet Lab members who provided a great atmosphere with their intellectual input on the bench side and at the meeting room. In order of appearance: Thibaut, Zuza, and Julian. In addition to the many wonderful project students. Finally, to my wonderful colleagues who made the South Coast shine brighter with their presence. As the cliché goes, there are too many to name but special mention to David, Tomisin, and Grant. Their idiosyncrasies ensured that I was never homesick.

My family had continued to provide support when times were most difficult and were not shy to let me know when they were proud of hard work and success but reminding me always to be humble. My mother and late-grandmother may have never understood what I was doing, but they were there when I needed.

It would be criminal to not extend my gratitude to both Professor Tony Carr and Professor Aidan for allowing me to pick myself up again. In both being able to complete what seemed like a Sisyphean task and providing professional support to step on to the next part of my life. Dr Kostas Thalassinou for not allowing me to take it easy with submission once I was back in a laboratory environment by continually pressuring me to complete the thesis. And of course, Dr Bilge Argunhan for helping to find the diamond I buried in the rough.

Lastly, a brave young man by the name of Mohamed Hadi who never despairs and has continually laughed at defeat's face. Truly, his chest houses a lion's heart.

"You never finish writing a book. You stop editing it."

Table of Contents

Acknowledgments	iii
Table of Figures	viii
Abbreviations	x
Chapter 1: Introduction.....	1
1.1. DNA and the Damage Response.....	1
Figure 1.1: DNA, RNA and Protein	3
1.2. DNA damage formation	5
1.2.1. Ultra-Violet Rays	5
1.2.2. High Energy Radiation (γ -rays and x-rays)	5
1.2.3. Free radicals.....	6
1.2.4. Perturbed biological processes.....	7
Figure 1.2: Map of sources of DNA Damage, Their Molecular Consequences and Possible Repair Pathways.....	9
1.3. Double Strand Break Repair Strategies	11
1.3.1. Double Strand Break	11
1.3.2. Non-Homologous End-Joining	12
1.3.3. Homologous Recombination	13
1.3.4. Microhomology-mediated end joining	14
Figure 1.3: Causes of DSB and Possible Repair Pathways.....	16
1.4. Histone Variants and Their Post-Translational Modifications	18
1.4.1. Introduction to HPTMs	18
Figure 1.4: DNA Packaging, Nucleosome and Histone Post-Translational Modifications	19
Figure 1.5: Post-translational modification of lysine, serine, arginine, and their mass changes.....	25
Figure 1.5: Post-translational modification of lysine, serine, arginine, and their mass changes (continued)	27
1.4.2. Histone Acetylation	29
1.4.3. Histone Methylation	31
1.4.4. Histone Phosphorylation	36
1.4.5. Histone Ubiquitination	38
1.4.6. Lysine Propionylation	40
Table 1.1: Chart of Characterised Histone Residues Susceptible to Acetylation, Methylation, Phosphorylation and, Ubiquitination and Histone Modifying Enzymes.....	42
1.5. Biological Mass Spectrometry	43
1.5.1. Introduction to mass spectrometry	43
Figure 1.6: Stages of Events within an LC-MS Orbitrap Mass Spectrometer	44
1.5.2. Ionization	46
Figure 1.7: Factors Determining Which Ionization Method To Use	49
1.5.3. HPLC separation vs direct infusion	51
Figure 1.8: Targeted LC-MS-MS	54
1.5.4. Orbitrap	56
Figure 1.9: MS/MS Compatible Instruments: Orbitrap vs Q-IMS-TOF.....	57
1.5.5. Tandem Mass Spectrometry: pseudo-Selected Reaction Monitoring.....	59
Figure 1.10: Tandem Mass Spectrometry (MS/MS) and Applications.....	60
1.5.6. LC-MS/MS compared to Immunoassays.....	62
Table 1.2: Comparison of the capabilities and feasibility immuno-assays and LC-MS	66
1.6. HPTMs in the DDR; An LC-MS/MS Challenge (preface to results)	67
Chapter 2: Methods and Materials.....	68

2.1. Culturing of Cells and cell lines used	68
2.2. DNA Damage: Ionising Radiation.....	68
2.3. Nuclear Extract	69
2.4. Acid Extraction of Histones	71
2.5. Micrococcal nuclease treatment	71
2.5.1. Agarose gel electrophoresis	73
2.6. Immunoprecipitation of γ H2AX.....	73
2.7. SDS-PAGE and Western Blot.....	74
2.8. Derivatization and digestion of peptides	76
2.8.1. In-solution derivatisation.....	76
2.8.2. In-gel derivatisation	76
2.8.3. In solution trypsin digestion of peptides	78
2.8.4. On-bead Trypsinisation.....	78
2.9. Peptide Desalting and Detergent Removal	79
2.9.1. Stage-Spin column desalting.....	79
2.9.2. Sera-Mag CM clean up of peptides.....	80
2.10. Phosphoenrichment.....	82
2.11. Nano-LC/MS	83
2.12. Data Analysis	84
Table 2.1 H2AX Control Peptides Used for Quantifications	85
Table 2.2 Control Peptides Used for Quantifications of pSRM	85
Table 2.3 Synthetic peptides and recombinant histones	85
Chapter 3: Global H2AX Analysis	87
3.1. Introduction	87
3.1.1. H2AX	87
3.1.2. S139 phosphorylation	89
3.1.3. H2AX Kinase	91
3.1.4. MDC1	91
3.1.5. 53BP1	92
3.1.6. Y142 phosphorylation.....	93
3.1.7. H2AX levels across cell lines and in the genome	94
3.2. Results	95
3.2.1. S139 phosphorylation	95
Figure 3.1. MS/MS Fragmentation of ATQASQEY S139 or Y142 phosphorylated precursors	97
Figure 3.2. Quantification of global γ H2AX S139 phosphorylation after DNA damage	100
3.2.2. Y142 phosphorylation.....	102
Figure 3.3. Y142 Phosphorylation	104
3.2.3. H2AX abundance	106
Figure 3.4. H2AX Abundance	107
Table 3.1. H2AX Abundance	109
Table 3.1. H2AX Abundance (Legend)	110
3.3. Discussion	111
1.1. Chapter 4: The Development and Optimisation of γH2AX-IP-MS Method....	115
4.1. Introduction	115
4.1.1. Rationale of Choice of DDR Proteins	116

Figure 4.1. Amplifying DDR-HPTM For Proteomic Screen	118
4.1.2. Potential Technical Issues	125
Figure 4.2 Common problems with antibodies	127
4.2. Results	134
4.2.1. 53BP1-3xFLAG Pull Down	134
Figure 4.3. 53BP1: 3xFLAG Pull Down	137
4.2.2. Lymphoblastoid for DNA Damage Proteomic Screen	144
Figure 4.4: Antibody Selection for γ H2AX IP	145
4.2.3. Native or Crosslinking for γ H2AX IP	148
4.2.4. Antibody Selection for γ H2AX IP	150
Figure 4.5: DOE Algorithm for antibody validation for the use of γ H2AX IP-MS	153
4.2.5 γ H2AX IP optimisation: λ -Phosphatase Treatment	155
Figure 4.6: γ H2AX IP optimisation: λ -Phosphatase Treatment	157
4.3.8 MEFS as control for antibody specificity	159
Figure 4.7: MEFS as control for antibody specificity	161
4.2.6 IP of γ H2AX with stringent conditions	164
Figure 4.8 IP of γ H2AX with stringent conditions	167
4.2.7 Antibody Titration	169
Figure 4.9: Antibody Titration	170
4.2.8 Mnase optimisation	172
Figure 4.10 Mnase Optimisation	174
Figure 4.11 Western blot of γ H2AX IP timecourse.	179
Chapter 5: LC-MS Quantification of HPTMs after IR-Induced DNA Damage	181
5.1 Introduction	181
5.1.1. H2AX Modifications	181
5.1.2. Ubiquitination of H2A(X)	182
5.1.3. H3K9 PTMs	182
5.1.4. Histone Docking Sites of 53BP1	183
5.1.5. H3K14ac	184
5.1.6. H3K36	185
5.1.7. H3K56ac	185
5.1.8. H4K16Ac	185
5.1.9. H2AZ	186
5.2. Results	187
5.2.1. Increase in H2AX S139 phosphorylation but not growth in enrichment of other variants of H2A	187
Figure 5.1. Increase in H2AX S139 phosphorylation but not increase in enrichment of other variants of H2A	189
5.2.2. Increase in Histone H3K9 Dimethylation	191
Figure 5.2. Increase in Histone H3K9 Dimethylation	194
5.2.3 There is no difference in Histone H3K27 methylation levels post DNA damage	196
Figure 5.3: No difference in Histone H3K27 methylation levels post DNA damage	198
5.2.4. There is no difference in Histone H3K36 di-/trimethylation levels post DNA damage	200
Figure 5.4 No difference in Histone H3K36 di-/trimethylation post DNA damage	202
5.2.5. There is a decrease in H3K79me1 at sites local to damage	204
Figure 5.5. Decrease in H3K79me1 at sites local to damage	206
5.2.6. There is a decrease in H4K20me1 at local sites of damage	208
Figure 5.6. Decrease in H4K20me1 at local sites of damage in irradiated LCLs	210
5.2.7. Small increase in Histone H4 Lysine acetylation post DNA damage	212
Figure 5.7. Small increase in Histone H4 Lysine acetylation post DNA damage	215
5.2.8. H3K14 acetylation does not change post-DNA damage at either local or global levels	219
Figure 5.8. H3K14 acetylation does not change post-DNA damage at either local or global levels	221
5.2.9. Acetylation of H3K18 increases at a local level following DNA damage	223

Figure 5.9. Acetylation of H3K18 increases at a local level following DNA damage.....	225
5.2.10. Acetylation of H3K23 increases at a local level following DNA damage.....	227
Figure 5.10. Acetylation of H3K23 increases at a local level following DNA damage.....	229
5.2.11. Decrease in H2AZ following ionizing radiation	231
Figure 5.11. Decrease in Histone H2AZ following ionizing radiation	233
5.2.12. Undetected Known DDR-related Histone post-translational modifications.....	235
Chapter 6: Concluding Discussion	243
6.1 Summary of Results	243
6.2 Design of Experiment Audit	245
6.2.1 LCLs: Use more cells and a lower dosage of γ ray	246
6.2.2. Exclusively Use S1; replenish S2 with more Mnase	248
6.2.3. Derivatisation.....	248
6.2.4 Controls to test region of the genome for Mnase activity	249
6.2.4.1. Acid Extraction	249
6.2.4.2. H2AX	249
6.2.4.3. Carrying out H4-IP-MS	250
6.3. Future Applications	251
6.3.1. H2AX Symmetry	251
6.3.2. Modular DOE	252
6.3.3. Late Repairing Breaks	253
Bibliography	254
Appendices	272
Appendix I: Table of enriched phosphorylated non-histone peptides. (Chapter 3)	272
Appendix II: Hatimy, A.A. et al., 2015. Histone H2AX Y142 phosphorylation is a low abundance modification. <i>International Journal of Mass Spectrometry</i> , 391, pp.1–7.	282

Table of Figures

Figure 1.1: DNA, RNA and Protein	3
Figure 1.2: Map of sources of DNA Damage, Their Molecular Consequences and Possible Repair Pathways	9
Figure 1.3: Causes of DSB and Possible Repair Pathways	16
Figure 1.4: DNA Packaging, Nucleosome and Histone Post-Translational Modifications	19
Figure 1.5: Post-translational modification of lysine, serine, arginine, and their mass changes	25
Figure 1.5: Post-translational modification of lysine, serine, arginine, and their mass changes (continued)	27
Table 1.1: Chart of Characterised Histone Residues Susceptible to Acetylation, Methylation, Phosphorylation and, Ubiquitination and Histone Modifying Enzymes	42
Figure 1.6: Stages of Events within an LC-MS Orbitrap Mass Spectrometer	44
Figure 1.7: Factors Determining Which Ionization Method To Use	49
Figure 1.8: Targeted LC-MS-MS	54
Figure 1.9: MS/MS Compatible Instruments: Orbitrap vs Q-IMS-TOF	57
Figure 1.10: Tandem Mass Spectrometry (MS/MS) and Applications	60
Table 1.2: Comparison of the capabilities and feasibility immuno-assays and LC-MS	66
Table 2.1 H2AX Control Peptides Used for Quantifications	85
Table 2.2 Control Peptides Used for Quantifications of pSRM	85
Table 2.3 Synthetic peptides and recombinant histones	85
Figure 3.1. MS/MS Fragmentation of ATQASQEY S139 or Y142 phosphorylated precursors	97
Figure 3.2. Quantification of global γH2AX S139 phosphorylation after DNA damage	100
Figure 3.3. Y142 Phosphorylation	104
Figure 3.4. H2AX Abundance	107
Table 3.1. H2AX Abundance	109
Table 3.1. H2AX Abundance (Legend)	110
Figure 4.1. Amplifying DDR-HPTM For Proteomic Screen	118
Figure 4.2 Common problems with antibodies	127
Figure 4.3. 53BP1: 3xFLAG Pull Down	137
Figure 4.4: Antibody Selection for γH2AX IP	145
Figure 4.5: DOE Algorithm for antibody validation for the use of γH2AX IP-MS	153
Figure 4.6: γH2AX IP optimisation: λ-Phosphatase Treatment	157
Figure 4.7: MEFS as control for antibody specificity	161
Figure 4.8 IP of γH2AX with stringent conditions	167
Figure 4.9: Antibody Titration	170
Figure 4.10 Mnase Optimisation	174
Figure 4.11 Western blot of γH2AX IP timecourse.	179
Figure 5.1. Increase in H2AX S139 phosphorylation but not increase in enrichment of other variants of H2A	189
Figure 5.2. Increase in Histone H3K9 Dimethylation	194
Figure 5.3: No difference in Histone H3K27 methylation levels post DNA damage	198
Figure 5.4 No difference in Histone H3K36 di-/trimethylation post DNA damage	202

Figure 5.5. Decrease in H3K79me1 at sites local to damage	206
Figure 5.6. Decrease in H4K20me1 at local sites of damage in irradiated LCLs	210
Figure 5.7. Small increase in Histone H4 Lysine acetylation post DNA damage	215
Figure 5.8. H3K14 acetylation does not change post-DNA damage at either local or global levels	221
Figure 5.9. Acetylation of H3K18 increases at a local level following DNA damage	225
Figure 5.10. Acetylation of H3K23 increases at a local level following DNA damage	229
Figure 5.11. Decrease in Histone H2AZ following ionizing radiation	233
Appendix I: Table of enriched phosphorylated non-histone peptides. (Chapter 3)	272

Abbreviations

-/-;	Knockout gene
+/+;	Intact Gene
53BP1;	p53-binding protein 1
Ab;	Antibody
Ac;	Acetylation
acetyl-	
CoA;	Acetyl coenzyme A
ACN;	Acetonitrile
AESF;	4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride
Anp32;	Acidic (leucine-rich) nuclear phosphoprotein 32 kDa
APCI;	Atmospheric Pressure Chemical Ionisation
APPI;	Atmospheric Pressure Photo Ionisation
APS;	Ammonium Persulphate
ATM;	Ataxia Telangiectasia mutated
ATR;	Ataxia Telangiectasia And Rad3-Related Protein
ATRX;	Alpha-thalassemia/mental retardation, X-linked
BAF;	Barrier-to-autointegration factor
BER;	Base excision repair;
BLM;	Bloom syndrome protein
BP;	Base pair
BRCA1;	Breast cancer 1, DNA repair associated
BRCA2;	Breast cancer 2, DNA repair associated
BRCT;	BRCA1 C Terminus
Cas9;	CRISPR-associated protein
C-Trap;	Curved Linear Trap
CAF-1;	Chromatin Assembly Factor 1
ChIP;	Chromatin immunoprecipitation
CI;	Chemical Ionisation
CID;	Collision Induced Dissociation
CK2;	Casein kinase 2
CMV;	Cytomegalovirus
CRISPR;	Clustered regularly interspaced short palindromic repeats
CtIP;	CtBP-interacting protein
Da;	Dalton
DDR;	DNA Damage Response
DMEM;	Dulbecco's Modified Eagle Medium
DMSO;	Dimethyl sulfoxide
DNA;	Deoxyribonucleic Acid
DNA-PKcs;	DNA-dependent protein kinase, catalytic subunit
DNMT3;	DNA Methyltransferase 3

DOE;	Design of Experiment
DOT1;	Disrupter of telomeric silencing
DSB;	Double-Strand Break
dsDNA;	double-stranded DNA
E. coli;	Escherichia coli
EBV;	Eppstein Barr Virus
ECL;	enhanced chemiluminescence
EDTA;	Ethylenediaminetetraacetic acid
EHz;	Exaherts
EI;	Electron Ionisation
ER- α ;	Oestrogen Receptor α
ESI;	Electrospray Ionisation
FA;	Formic Acid
FWHM;	Full width at half maximum
GCN5;	General Control Nonderepressible 5
Gy;	Gray
HAT;	Histone Acetyle Transferase
HCD;	Higher-energy collisional dissociation
HCT116;	Human colon cancer cell
HDAC;	Histone deacetylase
HEK293T;	Human embryonic kidney 293
HeLa;	Henrietta Lacks
HKMT;	Histone KMT
HP1;	Heterochromatin Protein 1
HPLC;	High Pressure Liquid Chromatography
HPTM;	Histone post-translational modification
HR;	Homologous recombination
HRR;	Homologous recombination repair
HU;	Hydroxyurea
HUS1;	Human homologs of Schizosaccharomyces pombe rad1
ICL;	interstrand crosslink;
IgG;	Immunoglobulin G
IR;	ionizing radiation;
Jmjc;	Jumonji-domain-containing
KAP1;	KRAB-associated protein-1
KAT;	Lysine Acetyl Transferase
KB;	Kilobases
KDMT;	Lysine Demethyl Transferase
KMT;	Lysine Methyl Transferase
kV;	kiloVolt
L3MBLT1;	Histone Methyl-Lysine Binding Protein
LC-	
MS/MS;	Liquid Chromatography-Tandem Mass Spectrometry

LCL;	Lymphoblastoid Cell Line
LSD1;	Lysine-specific histone demethylase 1A
LTQ;	Linear Trap Quadrupole
m/z;	mass / charge
MALDI;	Matrix Assisted Laser Desorption Ionisation
MB;	Megabases
MCM;	Minichromosome maintenance
MDC1;	Mediator of DNA damage checkpoint protein 1
Me;	Methylation
MeCN;	Acetonitrile
MEF;	Mouse Embryonic Fibroblasts
MeOH;	Methanol
MeV;	Mega Volt
MMC;	Mitomycin C;
MMS;	Methyl methanesulfonate
MMSET;	Multiple myeloma SET domain
Mnase;	Micrococcal nuclease
MOF;	Males-absent on the first p
MRE11;	Meiotic recombination 11
MRN;	MRE11-RAD50-NBS1
Mus81;	US81 endonuclease homolog (<i>S. cerevisiae</i>)
N3-meA;	N3-methyladenine
N7-meG;	N7-methylguanine
nanoLC;	Nano flow liquid chromatography
NBS1;	Nibrin
NEM;	N-Ethylmaleimide
NER;	Nucleotide excision repair;
NHEJ;	Non-homologous end joining;
NIB-250;	Nuclei-Isolation Buffer
NP40;	nonyl phenoxypolyethoxyethanol
NuRD;	Nucleosome Remodeling Deacetylase
PAB2;	Polyadenylate-binding protein 2
PARP-1;	Poly [ADP-ribose] polymerase 1
PARP-2;	Poly [ADP-ribose] polymerase 2
PCR;	Polymerase Chain Reaction
PHD;	Plant Homeo Domain
Phos;	Phosphorylation
PIKK;	Phosphatidylinositol 3-kinase-related kinase
PKC δ ;	protein kinase C- δ
PMSF;	phenylmethane sulfonyl fluoride
Pol δ ;	DNA Polymerase δ
Pol θ ;	DNA Polymerase θ
Pol λ ;	DNA Polymerase λ

Pol μ ;	DNA Polymerase μ
PRC 1;	Polycomb Repressor Complex 1
PRC 2;	Polycomb Repressor Complex 2
Prop;	Propionylated
pSRM;	pseudo-SRM
PTM;	Post-Translational Modifications
Q-TOF;	Quadrupole-Time of Flight
qPCR;	Quantative-PCR
QQQ;	Triple Quadrupole
Rad1;	Radiation Sensitive 1
Rad50;	Radiation Sensitive 50
Rad51;	Radiation Sensitive 51
Rad52;	Radiation Sensitive 52
Rad9;	Radiation Sensitive 9
RAG-1;	Recombination-Activating Gene 1
RAG-2;	Recombination-Activating Gene 1
RAP80;	Receptor-associated protein 80
RCF;	Relative Centrifugal Force
Rhp9;	Rad9 homologue in <i>S. pombe</i>
RING1B;	Ring Finger Protein 1
RMS;	Root Mean Square
RNA;	Ribonucleic Acid
RNA Pol II;	RNA Polymerase II
RNF168;	Ring Finger Protein 168
RNF8;	Ring Finger Protein 8
ROS;	reactive oxygen species;
RPA;	Replication Prtein A
RPM;	Revolutions per minute
RPMI;	Roswell Park Memorial Institute medium
S1;	Supernatant obtained after first spin down
S2;	Supernatant obtained after second spin down
SAM;	S-adenosyl methionine;
SDS;	Sodium dodecyl sulfate
SDSA;	Synthesis-dependent strand anealing
Ser139ph;	Phosphorylated Serine 139
SET;	Su(var)3-9, Enhancer-of-zeste and Trithorax
SILAC;	Stable Isotope Labeling by/with Amino acids in Cell culture
Sirt7;	Sirtuin 7
SMC5/6;	Structural maintenance of chromosomes 5/6
SRM;	Selected Reaction Monitoring
SSB;	Single-strand break
SSBR;	SSB repair;

ssDNA;	Single Stranded DNA
Suv39-1;	Su(var)3-9
SWI/SNF;	SWItch/Sucrose Non-Fermentable
TAE;	Tris base, acetic acid and EDTA
TALEN;	Transcription activator-like effector nuclease
TBS;	Tris-buffered saline
TBST;	TBS-Tween
TCA;	Trichloroacetic Acid
TdT;	Terminal deoxynucleotidyl transferase
TEAB;	Tetraethylammonium bromide
TEMED;	Tetramethylethylenediamine
TERT;	Telomerase reverse transcriptase
TFA;	Trifluoroacetic Acid
TFIID;	Transcription Factor II D
Tip60;	HIV-1 Tat interactive protein, 60kDa
TMZ;	Temozolomide
TOPO;	Topoisomerase
TRIM24;	Tripartite Motif Containing 24
TSS;	Transcription Start Site
U2-OS;	Human bone osteosarcoma cell line
UBC13;	Ubiquitin-conjugating enzyme E2 N
Ubi;	Ubiquitination
UV;	Ultraviolet.
UVB;	Ultra Violet B
WSTF;	Williams Syndrome Transcription Factor
XLF;	XRCC4-like factor
XRCC4;	X-ray repair cross-complementing protein 4
YEAT;	Yaf9, ENL, AF9, Taf14, Sas5 domain
βME;	2-Mercaptoethanol

Chapter 1: Introduction

1.1. DNA and the Damage Response

Deoxyribonucleic acid (DNA) is present in every living organism and is the set of instructions to create proteins, the building blocks of life. In brief, DNA is a double-stranded molecule that has each strand running anti-parallel to each other (WATSON & CRICK 1953). DNA is composed of four different nucleotides—guanine, adenine, thymine, and cytosine (**Figure 1.1a**). Genes of particular sequences of bases encode proteins. Critically, the two strands of DNA are held together via hydrogen bonding and the bases within a strand are bonded together via phosphodiester bonds. Information is stored in DNA using a triplet code: every three nucleotides within transcribed genes can be translated into one of twenty amino acid (Travers & Muskhelishvili 2015) (**Figure 1.1b**).

There are 64 different possible combinations that are used to code for one of twenty different amino acids (or a stop codon) during translation (**Figure 1.1c**). Therefore, DNA can be used to store information about which different protein sequences are in the repertoire of the cell i.e., which proteins the DNA encodes (Jackson, 2010). DNA is not only used for storing information; it can also be used to regulate gene expression depending on the specific promoter at the start of a gene; specific sequences of DNA can be used to initiate DNA replication; and the sequence can be repetitive in nature which gives it a structural function such as at the centromeres (Verdaasdonk & Bloom 2011) and telomeres (Lewis & Wuttke 2012).

Therefore, it is imperative that the sequence of DNA is maintained with high-fidelity: mutations in the sequence can impair gene regulation or protein structure and function. However, the maintenance of genome stability poses a unique challenge as the structure of DNA is under constant attack from both endogenous and exogenous sources. Below, I will briefly summarise the sources of DNA damage, the molecular consequences of this damage, how the cell defends itself from daily attacks, and the clinical consequences if the cell's defences are compromised.

Figure 1.1: DNA, RNA and Protein

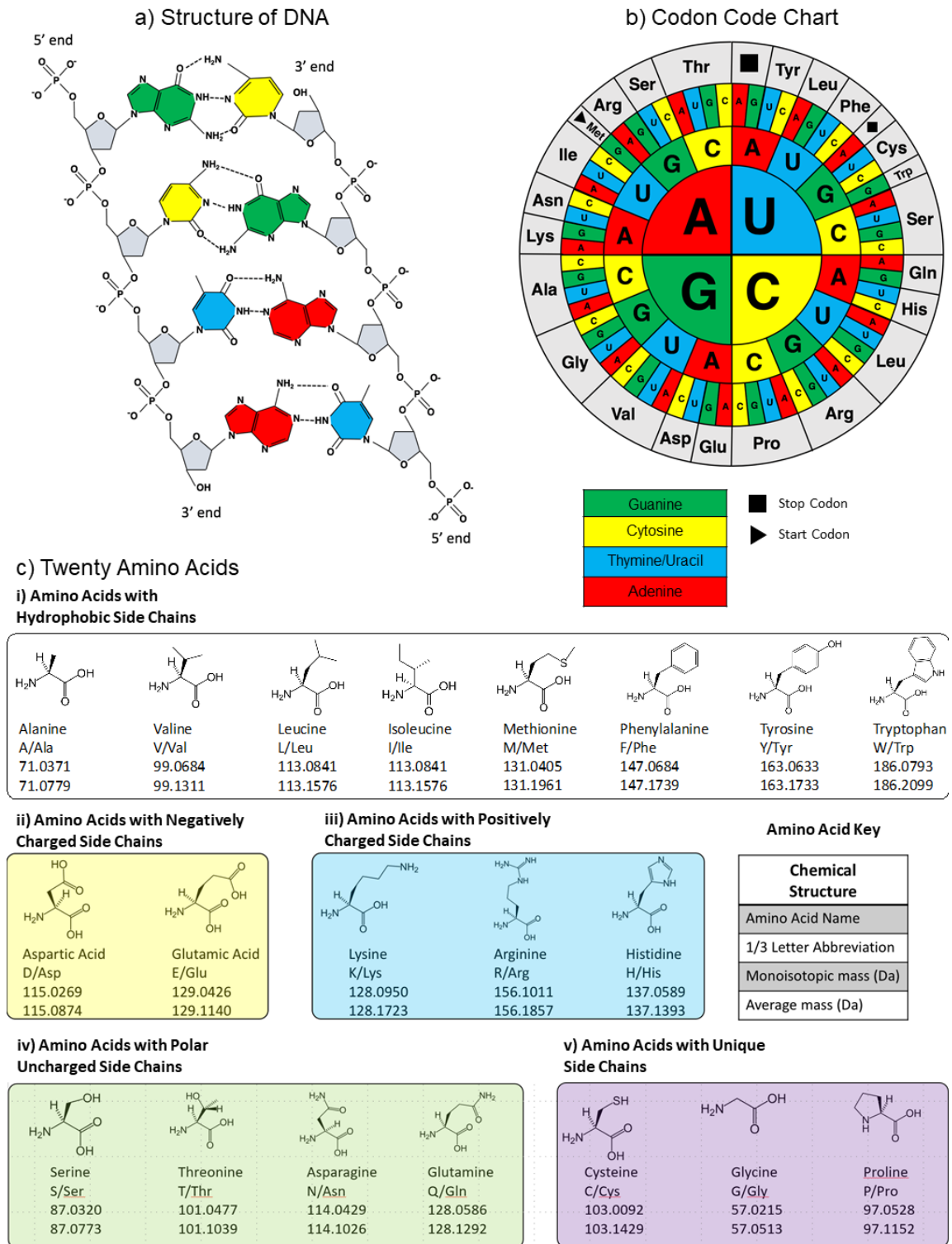


Figure 1.1 Legend: DNA, RNA and Protein

- a) **Structure of DNA:** Nucleotides are the most fundamental units of inheritance. Used both to transmit information and instruct cells. DNA is made up of four different nucleotides: adenine, guanine, cytosine and thymine. Guanine and adenine are purines, while cytosine and thymine are pyrimidines. Two DNA strands form a double helix and run anti-parallel to each other. The two strands bind together via hydrogen bonding. Adenine and thymine bond via two hydrogen bonds while cytosine and guanine use three hydrogen bonds. The sugar phosphate backbone gives DNA an overall negative charge.
- b) **Codon Code Chart:** During transcription (not shown) the DNA sequence is repeated in RNA. Thymine base is replaced with uracil. Three nucleotides in a sequence can either code for the start of translation (i.e. RNA to protein), for a particular amino acid or stop the translation altogether. Four bases picked three times leads to a maximum of 64 different combinations. However, there is redundancy because there are only twenty amino acid possibilities and a stop codon. Therefore, several amino acids can be encoded for by multiple codons.
- c) **Twenty Amino Acids:** The chemical properties of an amino acid is dependent on its R-group. As a result, the amino acids can be categorised according to the similarities of their properties. There are amino acids with hydrophobic side-chains; negatively charged side chains; positively charged side chains; uncharged side chains; and amino acids that cannot be categorised in the aforementioned groups.

1.2. DNA damage formation

1.2.1. Ultra-Violet Rays

As mentioned, DNA is continuously under attack from endogenous and exogenous agents i.e. sources of damage from within the cell and from the surrounding environment. One of the most well-known causes of DNA damage is UV radiation (Sinha & Hader 2002). Ultra-Violet B (UVB) light causes two adjacent thymine base pairs to bond to each other, forming pyrimidine dimers. If this aberrant structure is not resolved, a mutation can be introduced at the site of damage as the lesion can distort the structure of DNA by inserting bends, which prevents the replication or transcription machinery from being able to process the DNA at the site of the pyrimidine dimer (Sinha & Hader 2002) (**Figure 1.2**).

1.2.2. High Energy Radiation (γ -rays and x-rays)

Like UV and visible light, both γ -rays and x-rays are also on the electromagnetic spectrum but have much smaller wavelengths. Visible light has a wavelength ranging from approximately 400-700 nm compared to 10 pm (smaller than an atom) and 0.01-10 nm for γ -rays and x-rays, respectively. The shorter wavelengths lead to higher frequencies: 300 Exahertz (γ -rays) and 0.3-30 EHz (x-rays) compared to 300 THz (visible light). As a result of their higher frequencies, γ -rays and x-rays have higher energies compared to visible light: 1.24 MeV, 1.24-124 KeV and ~ 2.5 eV for γ -rays, x-rays, and visible light, respectively (Drake 2014). These higher energies are responsible for causing DNA damage that would not be possible with visible light or even UV. Such radiation may damage DNA through several means simultaneously: free radicals formed from nearby water molecules, which go on to react with the DNA molecule; directly ionising the DNA molecule by knocking-off electrons, leading to side reactions; or by disrupting the chemical bonds of the DNA molecule, such as hydrogen bonds or phosphodiester bonds.

If the noncovalent hydrogen bonds are broken, the strands of the DNA come apart (Vallur et al. 2002), but this is less deleterious compared to the phosphodiester bonds breaking. If the covalent phosphodiester bond is broken, it will result in the DNA molecule being physically broken. If only one of the two strands are broken, a single strand break (SSB) occurs and the cell can quickly repair the molecule by using the complementary strand as a template to synthesise a new strand and ligate the nicked DNA. However, if two phosphodiester bonds break on opposite strands in close proximity within the double helix, a double-strand break (DSB) is formed. The average dividing human cell would experience approximately 50 DSBs and 50,000 SSBs per day (Vilenchik & Knudson 2003; Caldecott 2008). Even though SSBs occur at a rate of three orders of magnitude higher, DSBs are the most cytotoxic form of DNA damage (Huertas 2010) (**Figure 1.2**).

1.2.3. Free radicals

Another source of DNA damage is oxidative damage from free radicals, which can either be from exogenous sources (via tobacco, pollutants, drugs, etc) or endogenous metabolism (Dizdaroglu & Jaruga 2012b). Free radicals are classed as reactive oxygen species and can react with double bonds within the structure of the DNA. These reactions can lead to adducts on the DNA molecule, which have to be removed for normal biological processes to proceed (Dizdaroglu & Jaruga 2012a). Peroxides (e.g. hydrogen peroxide, H_2O_2) are one such example of free radicals responsible for damage to the genome. Peroxides can form from one of two ways: radiolysis of intracellular water molecules by ionising radiation (IR), resulting in the formation of intermediates that then go on to form peroxides; or from normal cellular metabolism of oxygen (Sonntag 2006). In respiration, oxygen acts as an electron sink that is then reduced to form water. However, on rare occasions, the reduced oxygen molecule can be reduced to a

superoxide ($O_2^{\bullet-}$). If the superoxide leaks from the mitochondria into the rest of the cell, the superoxide can potentially be catalysed into peroxide (Sonntag 2006).

Another intrinsic free radical that can induce DNA damage is the hydroxyl radical (OH^{\bullet}), which is an intermediate in the radiolysis reaction from water to oxygen. Hydroxyl is also able to cause oxidative damage to DNA. However, whilst hydroxyl is a more reactive species compared to peroxide, it is also short-lived because of its very reactive nature (Gligorovski et al. 2015). In contrast, peroxide is able to remain stable long enough to enter the nucleus and get within close proximity to DNA, which makes peroxide more hazardous to the integrity of the genome compared to hydroxyl. To react with DNA, hydroxyl must be produced adjacent to the DNA molecule for it to have a chance to damage DNA as a consequence (Rogakou et al. 1998) (**Figure 1.2**).

1.2.4. Perturbed biological processes

When cells undergo processes that involve interacting with the DNA molecule directly, such as transcription or replication, it is not unusual for the cellular process to fail and damage the DNA as a consequence. DNA replication occurs during S-phase of the cell cycle and is initiated at specific sites known as replication origins (Podhorecka et al. 2010). The process of replication starting at the origin is termed origin firing and prior to origin firing, the replication proteins assemble themselves into a complex. It is here that the unwinding and synthesis of DNA takes place. During replication, it is possible for the replication fork to stall or collapse. A replication fork stalls when it encounters an impediment such as a damaged strand, resulting in the activity of the helicase and/or polymerase coming to an abrupt halt. The timeframe for the pause can be as long as it takes for the blockage to be removed (it is even possible for the pause to be maintained

indefinitely). Under conditions in which the replication fork persists for too long, a collapse can occur as a result of a single strand break being processed into a double strand break by endonucleases such as Mus81 (Gambus et al. 2009), (Labib & Hodgson 2007) & (Ciccio et al. 2003). Experimental approaches to initiate replication fork stalling use chemicals such as aphidicolin (DNA polymerase inhibitor) or hydroxyurea (HU, a ribonucleotide reductase inhibitor) (Heylmann & Kaina 2016).

Once the replication fork stalls, there is the potential for DNA damage signalling to occur. The minichromosome maintenance (MCM) helicase persists with unwinding DNA in the presence of stalled replication forks. The unwound DNA-primed for replication is single-stranded DNA (ssDNA)(Smith et al. 2009). Replication protein A (RPA) binds to ssDNA to protect it and prevent it from reannealing during homologous recombination. Similarly, RPA is recruited to stalled replication forks to bind to the ssDNA (Kano et al. 2006). RPA-bound ssDNA has the potential to induce the DNA damage response as it mimics the resected ends of a double strand break and activates ATR, as well as the Rad9-Rad1-Hus1 (9-1-1) complex. The recruitment of DNA damage-specific proteins to a replication environment demonstrates that damage to the DNA is an inevitable consequence of DNA metabolism and the cell has evolved an error-check system (**Figure 1.2**).

Figure 1.2: Map of sources of DNA Damage, Their Molecular Consequences and Possible Repair Pathways

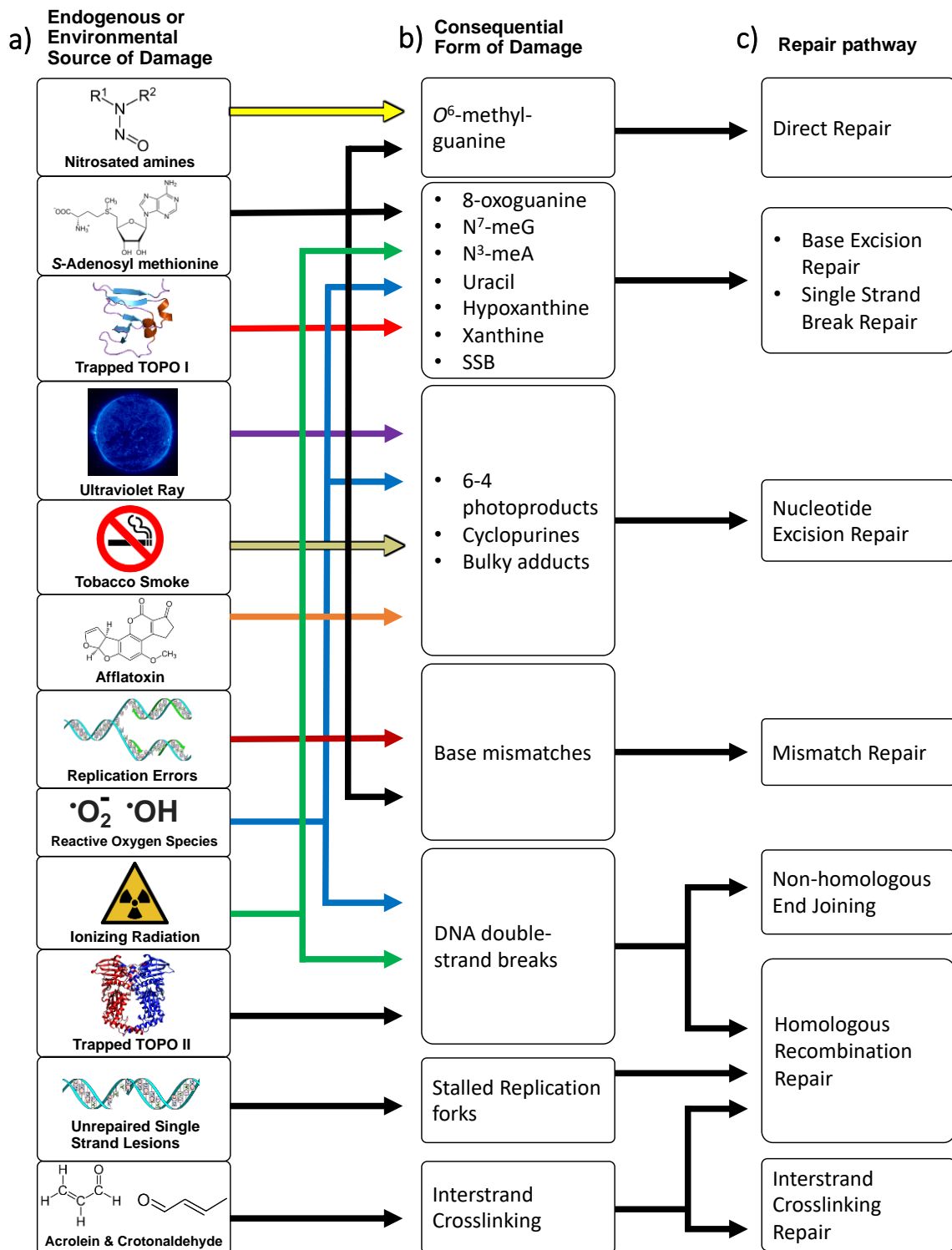


Figure 1.2 Legend: Map of sources of DNA Damage, Molecular Consequences and Possible Repair Pathways. *Adapted from* (Curtin 2012)

- a) **Left:** Endogenous and exogenous sources of DNA damage. Some DNA damaging agents can lead to multiple types of damage and have been colour coded to indicate possible outcomes e.g. Ionizing radiation (in green) can cause both DNA double-strand breaks, single-strand breaks and various adducts. As a result the multi-headed green arrow stems from the green box.
- b) **Middle:** Various forms of damages as a result of the exogenous and endogenous DNA damaging agents.
- c) **Right:** Organisms have evolved various and bespoke DNA Repair pathways to protect the genome from the different forms of DNA damage.

1.3. Double Strand Break Repair Strategies

1.3.1. Double Strand Break

In nature, there are various forms of damage repair pathways that take place. However, in this thesis, we will be focussing on Double Strand Breaks (DSB) and repair mechanisms pertaining to DSBs. A DSB is defined as a simultaneous break in both strands of the DNA molecule. DSBs can be induced by chemical agents, X-rays, γ-rays (**Figure 1.3a**), biological processes such as stalling of replication forks and induced by the cell during meiosis for meiotic recombination (**Figure 1.3b**). The reason DSBs are seen as the most hazardous form of DNA damage is due to the mutation and deletions that can arise as a result of the repair mechanisms that are recruited to repair the breaks. The mutations can then have downstream effects at both cellular and organismal levels. Moreover, DSBs can result in either acentric or dicentric chromosomes, which is where the chromosomes have either lost their centromere or fused with a fragment of another chromosome to gain a second centromere, respectively. Both scenarios can have devastating effects. During cell division, acentric fragments would result in a loss of genetic information as the chromosomes may not be distributed evenly between the daughter cells (Kaye et al. 2004). If fidelity of the repair mechanism has been compromised as a consequence of having lost genetic material, inter-chromosomal fusions (Robertsonian translocation) can take place, leading to genomic rearrangements, which are potentially carcinogenic if not dealt with by one of many cellular mechanisms evolved to deal with such an occurrence. In humans, Robertsonian translocations can potentially result in Down Syndrome (Celeste et al. 2002). If the DSB cannot be repaired, then the cell may induce apoptosis. At times, the breaks are deliberate and the exchange of genetic information is encouraged, but this programmed phenomenon is tightly regulated and temporally confined to a specific point

in meiosis (Celeste, Difilippantonio, et al. 2003) and VDJ Recombination. Repair pathways depend on which stage of the cell cycle the cell is in. DSB repair in mammalian systems is a complex process and the molecular mechanisms of the proteins involved are not well understood. Below, I summarise what is known about DSB repair in mammalian systems.

1.3.2. Non-Homologous End-Joining

The vast majority of DSBs in cycling cells are repaired by Non-Homologous End-Joining (NHEJ). NHEJ is most prominent in G1 and G2. In G1, NHEJ repairs nearly all DSBs and in G2, >80% of DSBs are repaired via NHEJ (Petersen et al. 2001). In terms of maintaining genomic fidelity, NHEJ can be either quick and efficient for 80% of breaks induced by IR where no resection takes place, or it can result in genetic deletion due to resection being activated. 53BP1 is believed to promote NHEJ over HR as the DSB repair pathway. (Panier & Boulton, 2014). NHEJ can occur during G1 where there is no template to copy a sequence from, unlike HR which requires a template to copy from (discussed below). The initial step of NHEJ is the recruitment of the heterodimer Ku70/Ku80 which will bind to the DSB to inhibit resection (Mimori & Hardin, 1986). The DNA-PKcs, belonging to the PIKK class of kinases, is recruited by Ku70/80 to the site of damage (Bassing et al. 2002). DNA-PKcs facilitates the interaction of Ku70/80 with XRCC4, Ligase IV and XLF. The BRCT domain of Ligase IV interacts with Ku70/80 only when Ku is a dimer. The interaction is catalysed by DNA-PKcs (Bassing et al. 2002). As a result, the interaction of XRCC4 with LigIV is then activated directly. Lig IV docks with XRCC4 via two BRCT domains that are present in the C-terminal region of DNA ligase IV. XLF then binds to both Ligase IV and XRCC4 to form a DNA ligation complex, which will ligate the broken DNA strands together (Petersen et al. 2001). Although this method of repairing DSBs is quick and efficient for most cases of NHEJ mediated repair of DSBs, approximately 20% of DSBs that are induced

by IR require a more intricate method of NHEJ repair (Riballo et al. 2004). A DSB may require further processing depending on the complexity of the situation. DSBs in heterochromatin may require additional processing. Therefore, more DDR response proteins involved in NHEJ will be recruited to the site (Riballo et al. 2004). If the site of damage is located in heterochromatic region and the cell is in G1 then the break will take longer to repair (Helt et al. 2005) (**Figure 1.3c**).

In addition to facilitating repair of DSBs via NHEJ, ATM and other repair proteins have been implicated in VDJ recombination. VDJ recombination is required for the generation of antigen receptor diversity and always occurs during the formation of new lymphocytes (Lou et al. 2006). RAG recombinase initiates VDJ recombination by binding and cleaving at specific sites known as recombination signal sequences that are located on either side of the V, D and J gene sections. Upon initiation of VDJ recombination by RAG recombinase, which induces a DSB, the RAG protein works in tandem with NHEJ repair machinery to ligate the newly recombined region together.

1.3.3. Homologous Recombination

By definition, homologous recombination-mediated repair can only take place when there is a sister chromatid present to copy a new sequence from. Thus, HR is discouraged in G1 as there is no sister chromatid. HR-mediated repair predominantly occurs during S-phase and G2 in the cell cycle. HR is a central component of meiosis, where instead of sister chromatids being used as a template, strand invasion between homologous chromosomes takes place (Kobayashi et al. 2004). In addition to being used for generating genetic diversity in meiosis, HR is also important for the repair of DSBs generated as a result of collapsed replication forks (Lou et al. 2006). HR is thought to be

a higher fidelity repair mechanism than NHEJ. However, it is only used in 20% of DSBs generated by IR in G2-phase of cells (Stewart et al. 2003). For HR to occur, single-stranded DNA is required. The single-stranded DNA can be generated from 5' resection during a DSB (Goldberg et al. 2003) or from a processed replication fork that collapsed. Replication Protein A (RPA) then coats the ssDNA. The Rad52/BRCA2 family of proteins and Rad51 paralogues then displace RPA and assist in the formation of the Rad51 nucleoprotein filament (Sung JBC 2003). Rad51 is the central protein in HR and serves two functions. Firstly, Rad51 conducts a homology search to identify a sequence similar to the 3' overhang. Secondly, upon finding a DNA strand with high sequence identity, which is normally a sister chromatid, the Rad51 filament invades the intact duplex DNA and pairs with the complementary strand (Stucki et al. 2005). As a result, the non-complementary strand, whose sequence is identical to the DNA comprising the Rad51 nucleoprotein filament, is displaced, forming a displacement loop (D-loop). Pol δ then utilises the 3' overhang as a primer and extends it. Depending on the processivity of this extension, a Holliday junction may or may not form. Importantly, the completion of HR without a Holliday junction always results in non-crossover outcomes, whereas the formation and subsequent resolution of a Holliday junction can result in either a crossover or non-crossover (B. Wang & Elledge 2007; Mattioli et al. 2012) (**Figure 1.3e**).

1.3.4. Microhomology-mediated end joining

Microhomology-mediated end joining (MMEJ), also known as alternative NHEJ, is a secondary method for repairing double-strand breaks via end-joining. It is believed that MMEJ takes over repair when regular NHEJ is compromised (Lee & Paull 2007). Exactly how MMEJ is chosen currently remains unknown. MMEJ is highly error-prone as its *modus operandi* results in deletions of biological information from the DNA sequence as

the pathway tends to delete approximately 5-25 base pairs flanking the site of damage in search of small tracts of homology (Ward et al. 2003). MMEJ's high propensity for error stems from the crude method in which the pathway will search for homology upstream and downstream of the relevant damaged strand. Upon finding 5-25 complementary base pairs on both strands, the strands will be aligned. Inevitably, there will be overhangs and mismatched bases, which are excised and removed respectively, resulting in deletion of the DNA in the intervening region between the repeats (Kleiner et al. 2015). In recent years, proteins implicated in other pathways have been shown to be involved in MMEJ such as ATM (Baldock et al. 2015) and 53BP1 (Baldock et al. 2015) (**Figure 1.3d**).

Figure 1.3: Causes of DSB and Possible Repair Pathways

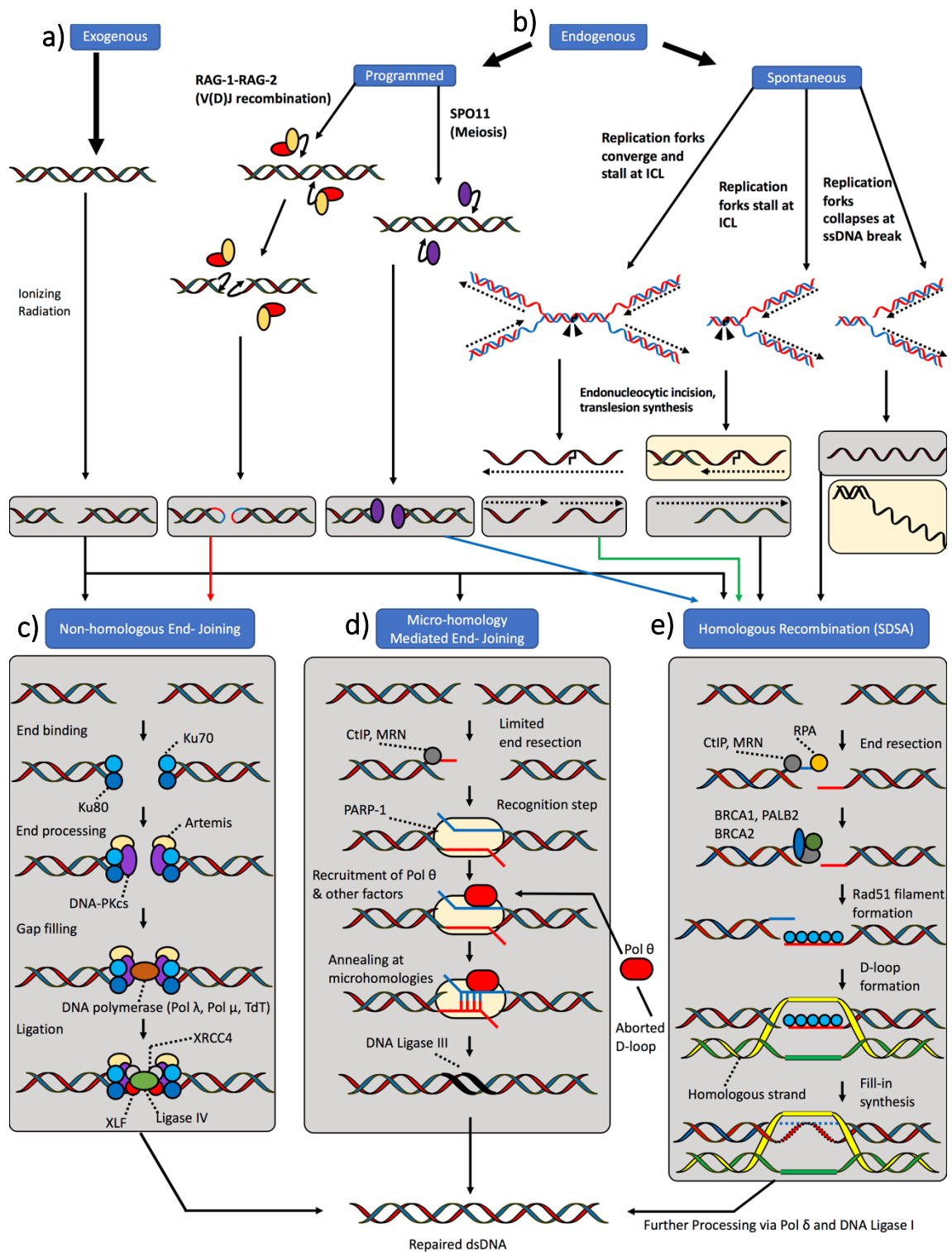


Figure 1.3 Legend: Causes of DSB and Possible Repair Pathways. *Adapted from:* (Ceccaldi et al. 2016) and (Chapman et al. 2012)

- a) **Exogenous causes of DSBs:** Ionising radiation and X-Ray can lead to DSBs. Most DSBs are repaired via NHEJ. However a proportion of DSB in G2 are repaired via HR and MMEJ
- b) **Endogenous causes of DSBs:** (i) Organisms have evolved auto-induction of DSBs to increase viability. During meiosis DSBs are induced for cross-over events to increase gene-pool variability. Also, developing lymphocytes will generate DSBs to randomly assemble different sections of genes to increase the different combinations of antigen receptors; thus reinforcing the immune system's defences. (ii) However, not all endogenous damage is scheduled. DSBs can be generated as a result of endogenous spontaneous damage such as ROS (not shown) or stalling of DNA processing such as during replication.
- c) **Repair via Non-Homologous End Joining (abridged):** Ku70 and Ku80 are recruited for end binding. Followed by Artemis and DNA-PKcs for End Processing. Gap filling carried out by DNA Pol λ and Pol μ . Ligation is completed by DNA Ligase IV.
- d) **Repair via Micro-Homology Mediated End-Joining (abridged):** CtIP and MRN induce limited end resection followed by PARP 1 for recognition of the lesion. Pol θ and other factors are recruited to complete annealing and gap filling. Ligation carried out by DNA Ligase III.
- e) **Repair via Synthesis-dependent strand annealing (Abridged).** There are several pathways with in HRR. SDSA is one form. CtIP and MRN induce end resection. RPA protects single strand overhangs. BRCA1, BRCA2, and PALB2 recruit Rad51. Rad51 filament formation induces homology searching. Upon D-loop formation, Pol δ and DNA Ligase I complete fill-in synthesis and ligation of the two ends, respectively. In some circumstances such as compromised HR pathway in cells, the D-loop may be aborted, resulting in recruitment of Pol θ where lesion is processed by the MMEJ pathway.

1.4. Histone Variants and Their Post-Translational Modifications

1.4.1. Introduction to HPTMs

DNA within the nucleus of eukaryotes is packaged within a molecular complex known as chromatin. The nucleosome, the principal component of chromatin, is comprised of an octamer that contains two copies of each of the four core histone proteins: H2A, H2B, H3, and H4. 147 base pairs of DNA is wrapped around a single octamer, which amounts to a strand of DNA encircling the octamer approximately 1.67 times (Goodarzi et al. 2008) (**Figure 1.4b&c**). When observed under a microscope, chromatin fibres have been documented to take shape in low ionic-strength environments. The appearance of the chromatin fibre has been described as "beads on a string". Chromatin fibre is the first level of chromatin compaction. A structure known as a solenoid (**Figure 1.4e**), which is a 30 nm fibre, can be observed when the ionic-strength is increased, which constitutes the second level of compaction (Kouzarides 2007) (**Figure 1.4f**).

The core histones are primarily globular. However, their N-terminals "tails" do not form a stable 3D structure but instead protrude out from the core. DNA binds to the nucleosome octomer through its minor groove, which has an electrostatic affinity for arginine residues (West et al. 2012). Many of the amino-acid residues of the N-terminal tails are receptive to a wide array of post-translational modifications (PTMS), including acetylation, methylation, and phosphorylation (Bannister & Kouzarides 2011) (**Figure 1.4d**).

Figure 1.4: DNA Packaging, Nucleosome and Histone Post-Translational Modifications

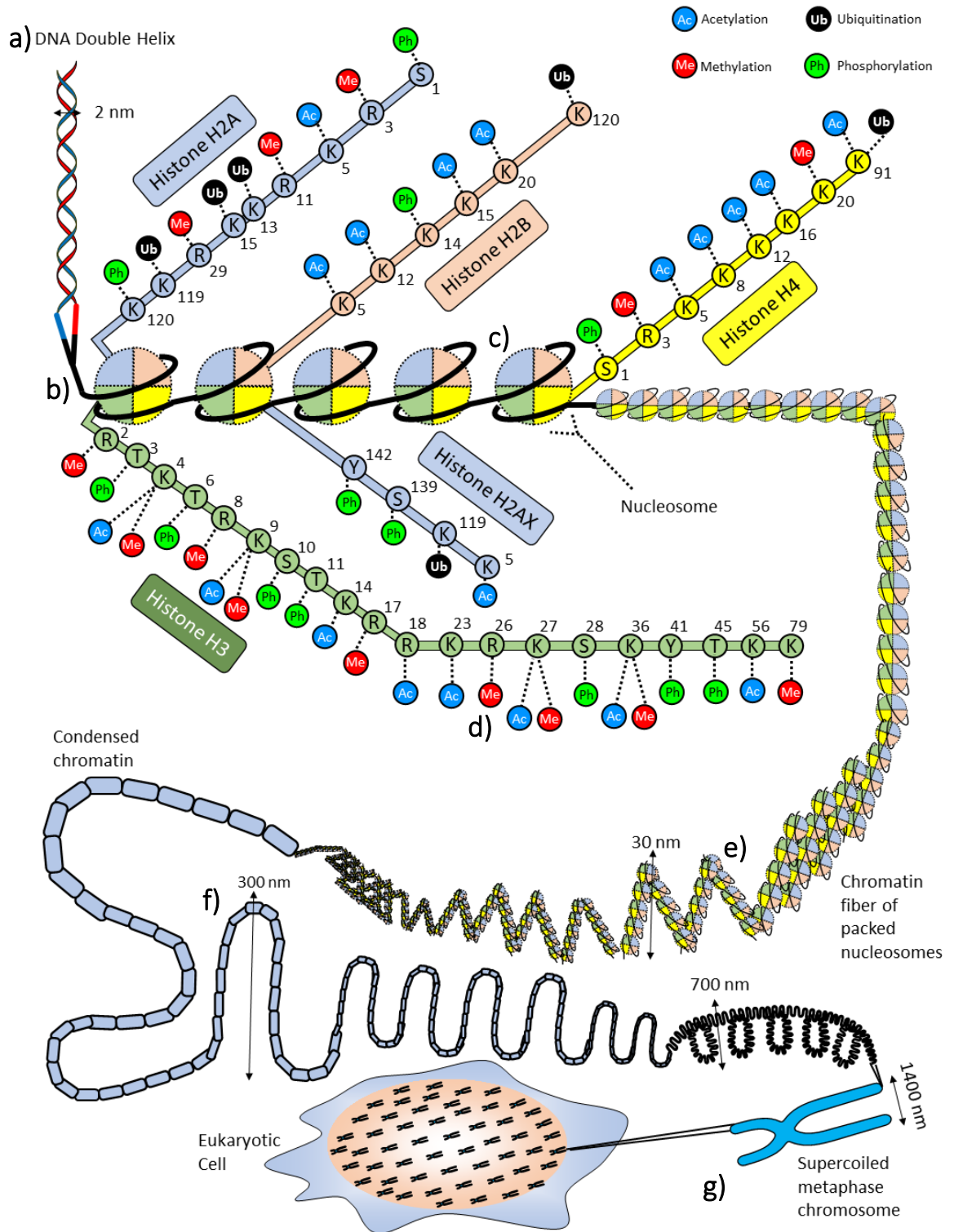


Figure 1.4 Legend: DNA Packaging, Nucleosome and Histone Post-Translational Modifications

- a) DNA Double Helix: the most simplest form of chromatin
- b) Nucleosome: 147 base pairs of DNA wrap around a histone octomer 1.67 times to form a nucleosome. This is recognised as the first level of chromatin compaction.
- c) Histone Octomer: Each octomer is comprised of two copies of each of the four histone variants: H2A, H2B, H3 and H4. These four histones are known as the core histones. In some organisms, Histone H1 (not shown) will seal the nucleosome
- d) Histone Post-Translational Modifications: Each histone has a terminus protruding from the octomer globular. These histone tails are subject to post-translational modifications. Highlighted are well-characterised HPTM of the core histones. Shown are lysine methylation, lysine acetylation, lysine ubiquitination, arginine methylation, and serine phosphorylation
- e) Second level of chromatin compaction: Nucleosomes fold to form solenoids. Solenoids fold further to forms 30 nm fibre
- f) Third level of chromatin compaction: 30 nm chromatin fibre folds further every 300 nm, which in turn forms 700 nm coils.
- g) Final level of chromatin compaction: a supercoiled chromosome found during metaphase in a eukaryotic cell

To illustrate their importance in maintaining genome stability, it is necessary to briefly cover the evolutionary history of histones. Histones are found in all eukaryotes and even some archaeobacteria (Eirín-López et al. 2009). Even though eubacteria have been observed to have a form of chromatin compaction, it is different to the system used by archaeobacteria and eukaryotes. *Escherichia coli*, which has been studied extensively, uses supercoiling for compaction; as its genome is only 4.6 Mega bases, this level of compaction is adequate (Sandman & Reeve 2000). Euryarchaea have been characterised to possess proto-histones that have the same function of wrapping DNA but they were missing eukaryotic features such as unfolded N-terminal and C-terminal tails (Arents & Moudrianakis 1995) and are wrapped by a shorter amount of DNA: 60 bp as opposed to 147 bp. This would imply that histone proteins evolved very early on as genome complexity grew with time. The importance of histones to eukaryotes in regards to organising a large genome can be properly highlighted when examining their genomes. As eukaryotes have grown more complex and become multicellular, their genomes have also had to increase in complexity. to illustrate this complexity, an *E. coli* cell has 4.6 million base pairs, whereas a haploid human cell possesses approximately 3 billion base pairs, which is several magnitudes larger. The human genome also contains approximately 20,000 genes (Lander et al. 2001) compared to an *E. coli* which only has about 4000 genes (Blattner et al. 1997). To assist with organising the genome, complex eukaryotes have not only evolved to increase the number of histone variants, which as this chapter will highlight have specialised roles, but have also increased the copy number of the same variants. Yeast, a simple eukaryote, possesses two copies of each of the four core histones. On the other hand, human Histone H4 alone is encoded by 14 genes

(The UniProt Consortium, 2017). The rise in copy number of histone genes can be partly attributed to whole genome duplication events (Malik and Henikoff 2003).

After a whole genome duplication event, evolutionary selective pressure would probably steer towards the genome keeping multiple copies of the histones for one of two reasons: it is better to have redundancy for crucial proteins, and with an increased genome size, it is necessary for the cell to have higher levels of histone proteins to protect the newly synthesised DNA. Through compaction, the DNA is protected from biological machinery that may process the DNA resulting in deletion or incorrect expression of a gene. For this same reason, as newly synthesised DNA strands emerge after DNA replication, the demand for histones would more than double immediately as the newly synthesised DNA would need to be protected and the cell needs to ensure that supply can keep up with demand. Possessing multiple copies of histone genes also allows for the emergence of histone variants. There are 15 genes that encode for Histone H4, which has no variants. Histone H3 has three variants encoded by 12 genes (The UniProt Consortium 2017) & (Marzluff et al. 2002). However, H2A and H2B have many different variants. Recent evidence also suggests that there is a proportional relationship between the size of the genome and the number of arginine residues present in its histone H2A variants, whereas an inversely proportional relationship has been observed between genome size and serine/tyrosine residues (Macadangdang et al. 2014). Histone variants can serve different functions. For instance, Histone H2AZ has been implicated in several biological contexts: chromatin compaction (Fan et al. 2002), the DNA Damage Response (DDR)(Y. Xu et al. 2012), meiosis (Yamada et al. 2017), and demarcating transcriptional start sites and gene bodies (Gervais & Gaudreau 2009) (Coleman-Derr & Zilberman 2012). Another variant,

H2AX, has been implicated in the DDR (Rogakou & Bonner 2000) and development (Cowell et al. 2007). The functions of the different variants will be discussed in detail further on. Selective pressures on histone function have resulted not only in them being highly conserved in their sequences but also in their PTMs. In mammals, it was found that not only was the function conserved at a molecular level but there was conservation for which cell types are regulated by the presence of a PTM (Baldeyron et al. 2011).

There are many sites at which histones can be modified and the number of known sites may be expected to increase with time (**Table 1.1**). In addition to the sites, the different types of modifications add complexity to histone PTMs. For example, lysine can be methylated in one of three ways: monomethylation, dimethylation and trimethylation, whilst arginine can also be methylated or dimethylated, asymmetrically or symmetrically i.e. where S-Adenosyl methionine (SAM) I dimethylates one amine exclusively in the R-group or where SAM II methylates methylates the two terminal amines in the R-group, respectively (**Figure 1.5 J-M**). Functionally, the histone PTMs can be divided into two categories: demarcation of global chromatin environments and managing DNA-protein based biological processes

The DNA wrapped around histones can become inaccessible which can hinder molecular functions such as transcription and repair. Some PTMs of histones weaken the interaction with DNA, which then allows access to biological machinery. Modifications can affect the structure of chromatin through one of two ways: perturbing the compaction between nucleosomes, which reveals chromatin, or by recruiting non-histone proteins. The loosely packed chromatin is known as euchromatin and the tightly bound chromatin is known as

heterochromatin (Bannister & Kouzarides 2011). A large number of different modifications allow for numerous possible functional responses. Even though there are numerous histone PTMs, in the context of the DNA damage response, the PTMs that we will be focussing on are acetylation, methylation, and ubiquitination of lysine residues, and phosphorylation of serine, threonine and tyrosine residues. Different domains are capable of recognising different PTMs. PHD and ATRX-DNMT3-DNMT3L recognise unmodified lysines, while Bromodomain and YEATs bind to acetylated lysines. There is a larger group of domains that can recognise methylation. These include: chromodomains, Tudor and Bromo-adjacent.

Figure 1.5: Post-translational modification of lysine, serine, arginine, and their mass changes

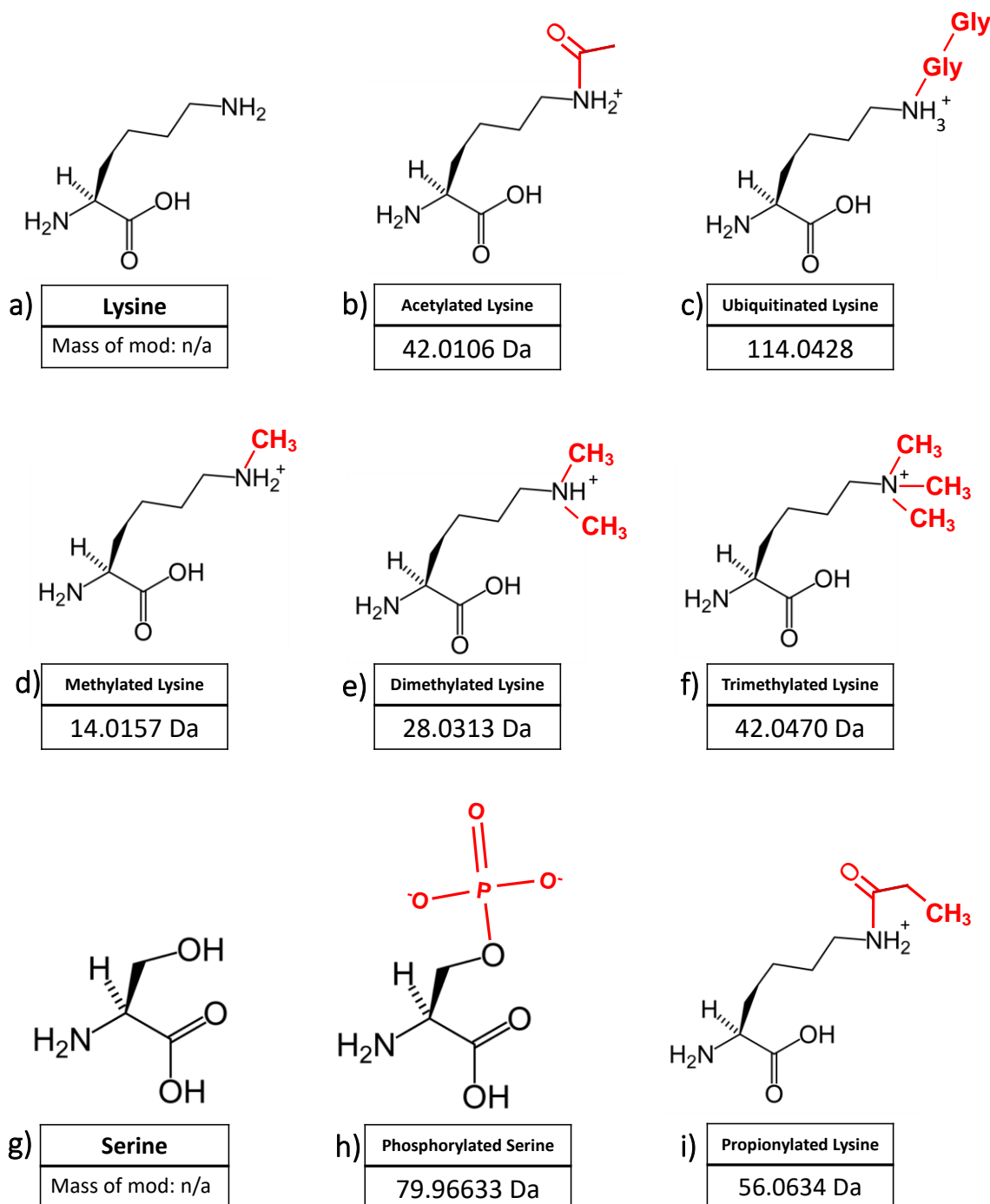


Figure 1.5 Legend: Post-translational modification of lysine, serine, arginine, and their mass changes

- a) Unmodified lysine: Nucleophilic residue. Susceptible to many PTMs
- b) Acetylated lysine: Δ mass of 42.0106 Da. Results in reduced overall positive charge. In regards to histones, results in opening up of chromatin. Usually results in transcriptional activation.
- c) Ubiquitinated lysine: Δ mass of 114.0428 Da. Involved in recruiting other factors e.g. during DDR H2AK15ubi acts as a docking site for 53BP1. Full ubiquitin not shown. Gly-Gly is artefact left covalently attached after tryptic digest
- d) Methylated lysine: Δ mass of 14.0157 Da. Associated with chromatin compaction but in certain cases can result in transcriptional activation e.g. H3K4me
- e) Dimethylated lysine: Δ mass of 28.0313 Da. Requires lysine monomethylation. Can be involved in transcriptional activation or repressed. Also involved in DDR e.g. H4K20me2.
- f) Trimethylated: Δ mass of 42.0470 Da. Requires lysine dimethylation. Can be involved in transcriptional activation or repression e.g. H3K36me3 and H3K9me3, respectively
- g) Unmodified serine: Highly polar residue.
- h) Phosphorylated serine: Δ mass of 79.96633 Da. Involved in signalling. Present in DDR e.g. H2AXS139phos and cell-cycle e.g. H3S10phos.
- i) Propionylated lysine: Δ mass of 42.0106 Da. In-vitro modification utilised to block unmodified histones to give uniform sized peptides after tryptic peptide. Presence is indicative of unmodified or mono-methylated lysines

Figure 1.5: Post-translational modification of lysine, serine, arginine, and their mass changes (continued)

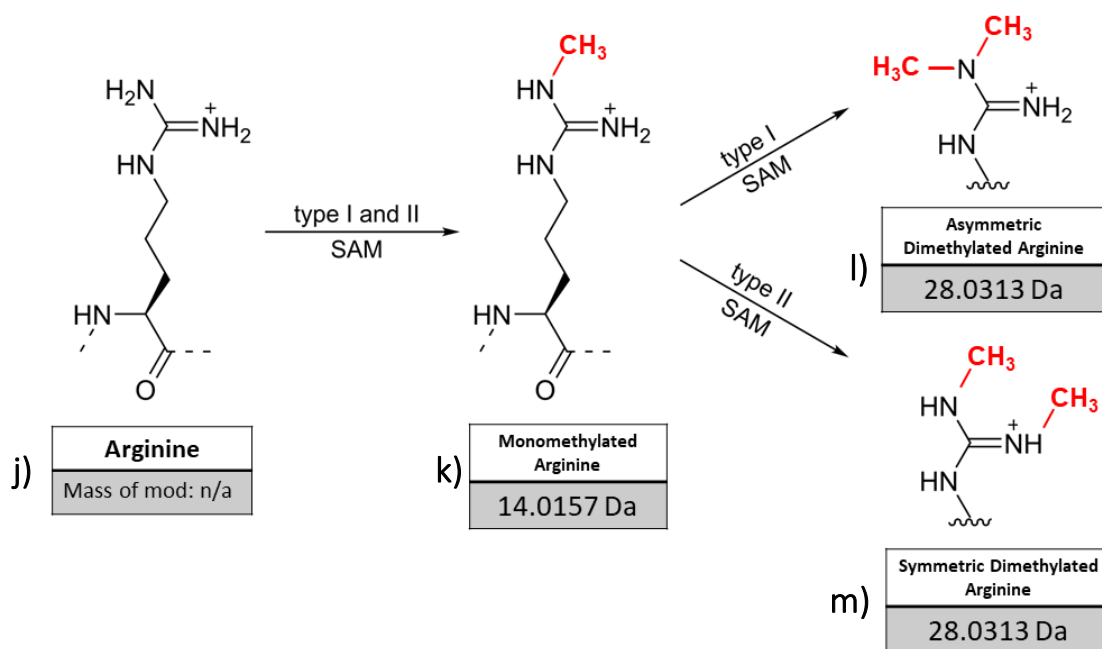


Figure 1.5 Legend: Post-translational modification of lysine, serine, arginine, and their mass changes (continued)

- j) Unmodified arginine: Nucleophilic residue. Susceptible to PTMs
- k) Monomethylated arginine: Δ mass of 14.0157 Da. In regards to histones, regarded as an intermediate PTM to further demethylation. Methylated by both type I and II S-Adenosyl methionine (SAM)
- l) Asymmetric dimethylated arginine: Δ mass of 28.0313 Da. Can act as a docking site for DNMT3A, a DNA methyltransferase, as seen on histone H4R3me₂. Methylated exclusively by type I SAM
- m) Symmetric dimethylated arginine: Δ mass of 28.0313 Da. Can block readers of adjacent residues carrying PTMs e.g. hinders access to H3K4me₃. Methylated exclusively by type II SAM

1.4.2. Histone Acetylation

Histone acetylation involves the addition of an acetyl group to lysine residues (**Figure 1.5b**). The acetyl group is composed of a methyl group bonded to a carbonyl group. Histone acetylation or deacetylation of lysine residues is known to regulate gene expression. Enzymes responsible for histone acetylation and histone deacetylation are known as histone acetyltransferases (HATs) and histone deacetylases (HDACs), respectively (Allfrey et al. 1964). HATs attach acetyl groups, which originate from acetyl-CoA, to lysines. HDACs carry out the opposite function of HATs by removing acetyl groups from lysines via hydrolysis. PTMs that alter the histone charge may change the compaction of chromatin (Kouzarides 2007). Histone acetylation and deacetylation are implicated in many biological processes such as nucleosome assembly, DNA replication, and transcription. HATs and HDACs work in an antagonizing fashion in the context of biological tasks that require chromatin to be relaxed. For example, HATs play a huge role during replication, but during transcriptional silencing, HDAC activity predominates.

There are two characterised families of HATs that add acetyl groups on to histone lysine residues. The variants known as Type-B HATs are responsible for transiently acetylating newly synthesised histones in the cytoplasm at sites H4K5 and H4K12. Once the histones have been deposited into the genome, the acetyl groups at these sites are removed almost instantaneously (Parthun 2007). There is more variety amongst the Type-A family of HATs, as Type-B family HATs share a lot of sequence homology within their group. Type-A HATs can then be categorised further (Hodawadekar & Marmorstein 2007).

Acetylation can affect chromatin in one of two ways. The overt way is through

neutralisation of the positive charge on lysine, resulting in the histone protein interacting less strongly with the negatively charged phosphate group of DNA (Grunstein 1997). It was hypothesised that HATs were involved in DNA transcription in some way (Luan et al. 2015). Transcription requires an open chromatin environment so that the machinery can get access to the DNA. Allis's group demonstrated that *Tetrahymena thermophila* possesses a gene that shares strong homology with GCN5 from yeast systems. Interaction of GCN5 with chromatin led to a reduction in chromatin compaction, making the DNA more accessible, and thus allowing the DNA to be available for biological processes such as transcription. These results directly demonstrated that acetylation and transcription occur concomitantly (Brownell et al. 1996).

A less obvious way in which acetylation can alter the chromatin environment is through the binding of bromodomain-containing proteins. Bromodomains are highly conserved regions of proteins that can be considered readers of lysine acetylation. Owen, et al. (2000) showed that the yeast HAT GCN5 possessed a bromodomain and was able to recognise various lysine residues from histone H4 that have been acetylated; it is not uncommon for proteins possessing bromodomains to have affinities for various acetylated lysine residues (Josling et al. 2012).

HDACs can counteract the activity of HATS. For instance, RNA polymerase needs to bind to the transcription start site so that other factors can be recruited downstream of it to initiate transcription. RNA pol II is required for the unwinding of the DNA double helix, otherwise processing of the template strand will not be possible (Pineda et al. 2015). However, nucleosomes act as hindrances to this and many other biological functions by physically blocking access to DNA. It is not uncommon for different HDACs to function

together in the same complex, which can result in more stringent removal of the acetylation (Miller et al. 2010).

1.4.3. Histone Methylation

Methylation involves the addition of a single or multiple methyl groups to an amino acid residue with a positively charged sidechain. Histidine, arginine and lysine can carry a maximum of one (monomethylation), two (dimethylation) or three (trimethylation) methyl groups, respectively (**Figure 1.5 d,e & f**). Even though methylation can take place at multiple amino acid residues (Greer et al. 2012) we will focus mostly on histone lysine methylation. Methylation can either promote transcription or suppress it. Complexity is added to its role since there are three possible modifications that can take place at the same lysine residue and multiple types of methyl-binding domains can distinguish these, allowing for the recruitment of a wide range of proteins.

Importantly, lysine methylation and acetylation are mutually exclusive; they cannot occupy the same lysine residue simultaneously. Methylation works differently to lysine acetylation as methylation does not neutralise the positive charge as it does not possess a polar carbonyl component. Enzymes that are responsible for methylating lysines are known as lysine methyltransferases (KMT). Lysines can only be methylated and demethylated in a sequential fashion, which means that a methyl group is either added or removed one at a time. For example, a lysine residue cannot be trimethylated without prior mono- and dimethylation. Over thirty KMTs have been described (Black et al. 2012) and they mostly possess the catalytic SET domain. Evidence of the first lysine demethylases (KDMs) was first shown by Shi, et al. (Y. Shi et al. 2004) with the discovery of Lysine-specific histone demethylase 1A (LSD1). In a contextual manner, some methylation can be transient such as those responsible for temporary compaction of the

chromatin during cell differentiation after stem cell generation, while other methylation sites can be permanent to the point that they are inherited through meiosis and mitosis, effectively forming epigenetic sites. As a result of the different methylation sites, some post-translational modifications will have faster turnover compared to others.

Methylation can be deposited by one of three families of proteins known as histone methyltransferases (KMT). Methyl groups that are deposited on histone lysines are sourced from S-adenosylmethionine. Two families of KMTs have been found to catalyse the methyl deposition on lysine residues: SET domain-containing proteins (Rea et al. 2000) and DOT1 (Feng et al. 2002). There are also two corresponding groups of proteins that carry out the opposite function of removing the methyl groups from lysines: amine oxidases such as LSD1 (Y. Shi et al. 2004) and JmjC domain containing dioxygenases that are iron-dependent (Tsukada et al. 2006). Both the JmjC-domain-containing KDMT and amine-oxidases share homology in both function and sequence from yeast up to mammalian systems. Some biological processes may require a permanent state of methylation such as structural regions, e.g. centromeres and telomere, or mitotically inherited silenced regions in differentiated cells that need to maintain uniformity of instructions with regards to gene expression and silencing. However, other regions may require sites of methylation to be malleable, such as: when cells are in the process of differentiating, which requires gene expression to differ from one cell type to another; or when cells are responding to environmental selective pressures including nutrient uptake and DNA damage. These factors result in there being large discrepancies between lysine residues for the duration of methylation turnover (Broering et al. 2015).

Presently, six lysine methylation sites have been characterised in detail: H3K4, H3K9,

H3K27, H3K36, H3K79 and H4K20 (Black et al. 2012). Generally, methylation at H3K4, H3K36 and H3K79 have been related to activation of transcription (Martin & Y. Zhang 2005), whilst H3K9, H3K27, and H4K20 have been implicated in repressing transcription. The presence of di- and trimethylation of H3K9 and H3K27 are indicative of heterochromatin and gene silencing (Kouzarides 2007). Both sites are susceptible to acetylation (Woo & W. H. Li 2012), so a rudimentary function would be to block lysines from being acetylated as both PTMs cannot occupy the same residue. Both sites being methylated would also function as signalling markers. When H3K9 has been methylated by Suv39-1, Heterochromatin Protein 1 (HP1) can dock on the methylated site, which initiates compaction by physically bringing the nucleosomes closer together via self-aggregation as they have an affinity for chromo-shadow domains. HP1 has also been shown to be present when deacetylation takes place (Saksouk et al. 2015). H3K9 methylation is found at sites of constitutive heterochromatin such as pericentromeric regions. Human Y chromosomes contain large regions of constitutive heterochromatin that are peppered with small regions that encode active genes (Quintana-Murci & Fellous 2001). Histone H3K27 trimethylation is catalysed by EZH2 which is part of the Polycomb Repressive Complex 2. Both H2AK119Ub and H3K27 methylation can be found at sites of facultative heterochromatin such as Barr bodies. The presence of methylation of H3K4 is linked with euchromatin, and specifically with genes that are active or fated for expression (Bernstein et al. 2002). H3K4me1 has been found to be present at enhancer sites across the genome (Heintzman et al. 2007). In contrast, H3K4me3 was found to associate much closer to the site of transcription by occupying regions corresponding to gene promoters (Santos-Rosa et al. 2002). It is intriguing that the same amino acid residue possesses two highly similar modifications, yet the localization and function of

each modification varies, despite the same common goal (i.e., upregulation of transcription). H3K4me1 is required for the chromatin remodeller BAF (human homologue to yeast SWI/SNF) to associate with histones. In vitro assays showed that the BAF45C subunit of BAF possesses a binding site for the single methylation but it will not bind to H3K3me3 (Local et al. 2018). However H3K4me3 was also shown to be required for the global recruitment of Transcription Factor II D (TFII D) through the TAF subunit of TFII D. TAF possesses a PHD binding domain which recognises H3K4me3 (Lauberth et al. 2013). In addition, H3K4me3 will recruit a different remodelling factor CHD1 which will loosen chromatin (Local et al. 2018)

As stated, H3K4 methylation is indicative of gene expression but that is not always the case. It is a modification that is also associated with different functions altogether in a different context. A modification possessing an antagonistic function to itself is not out of the realm of possibility, especially when the proteins it recruits can change in different environments. Both di and trimethylation of H3K4 have been documented in partaking in transcriptional repression (X. Shi et al. 2006). The mercurial function of the modification could be explained by cell cycle stage and/or DNA metabolism. Different effector proteins are expressed throughout the cell cycle that their presence at different times is akin to shift-workers on an assembly line. So if the region is scheduled to undergo silencing for a given biological function, the cellular support network can shift to accommodate the process by expressing effector proteins to silence the regions ad hoc. For example, ING2 repressor, which possesses a PHD domain that can read H3K4 methylation sites, will associate with HDAC at the same site (X. Shi et al. 2006). In another instance, H3K4me3 can be found in a bivalent domain i.e., associating with H3K27me3 at the same promoter,

in genes associated with development (Berstain et al 2006). The purpose of bivalency could be explained with being associated with development or rapid differentiation. In an environment where genes need to be quickly switched off or on it would serve as an advantage to have either option ready to go. H3K4me3 blocks recruitment of Polycomb Repressor Complex 2, which is involved in epigenetically silencing genes with KMT activity. In essence, the bivalent domain could be seen as fine-tuning of expression (Voigt et al. 2013).

Similarly to H3K4 methylation, H4K20 methylation is another example of a HPTM being involved in many cellular processes, which can also change with context: H4K20me1 is associated with transcriptional repression (Karachentsev 2005) and perturbed trimethylation H4K20 can have dire implications. Overexpression has been correlated with premature ageing of cells (Shumaker et al. 2006), whilst under expression has been linked to breast cancer (Tryndyak et al. 2006). H4K20me1/2 provides a binding site for 53BP1 at sites of DSBs. In cells with reduced H4K20me2, the cells have delayed 53BP1 focus formation, and thus lack an efficient repair system for DSBs (Panier & Boulton, 2014).

Methylation of the different residues gives an insight into how complex and intricate cellular instructions are and how much our understanding has progressed since simple Mendelian genetics. In turn, the complexities also demonstrate that our understanding of expression dynamics has shifted from binary (on/off) to a new paradigm of the cell's instructions being determined not only by the DNA sequence but the presence of DNA methylation, the histone post-translational modification and now which combination of HPTM is present at a given site. Theoretically, with each new condition, the number of

permutations of what can take place as a result of chromatin dynamics increases exponentially. However, the actual observed combinations are not as vast. The dynamics and permutations that are at play are vastly more intricate than any piece of software code and elucidating the cross-talk with systems biology will be an arduous task for years, if not decades. To conclude with methylation, it is not the modification itself that dictates what happens at a local level but the reader of the modifications that are present at any given moment, which have downstream effects on other chromatin-associated proteins.

1.4.4. Histone Phosphorylation

All of the core histones possess sites along their tails that are susceptible to phosphorylation by a number of protein kinases and can also be dephosphorylated by numerous phosphatases. Phosphorylation occurs at serine, threonine and tyrosine residues in histones (**Figure 1.5h**). It comprises the addition of a phosphate group to the hydroxyl in a serine or tyrosine residue. Phosphorylation is one of the most understood HPTMs; for decades, it has been recognised that kinases control signal transduction from the cell surface. In 1998, Rogakou, *et al.* (Rogakou et al. 1998) showed that serine 139 (128/129 in yeast) on H2AX is susceptible to phosphorylation following a double-strand break. Ser139ph is known as γ H2AX. The phosphatidylinositol 3-kinase-related kinase (PIKK) family, including ATM, ATR and DNA-PK, are responsible for the phosphorylation of S139 of H2AX (Harrison & Haber 2006). In humans, H2AX is the only H2A variant that possesses an elongated C-terminal tail, which includes S139.

Today, the presence of γ H2AX is routinely used for DSB assays as it has become synonymous with DNA damage. Phosphorylation of γ H2AX by the ATM kinase works as a signal for the DNA damage response (DDR) machinery. γ H2AX is detected quickly (<5 min)

after DNA damage and can spread across a large distance from the DSB: 1Mb in mammalian cells and up to 100kb in yeast cells. γ H2AX is responsible for recruitment of several DDR complexes to DSBs: Rhp9/53BP1 (Baldock et al. 2015); chromatin remodelers such as INO80; cohesin (Caron et al. 2012); and SMC5/6 (Nakamura et al. 2004). Without γ H2AX/ γ H2A, it has been shown that in both yeast and mammalian cells, the DNA damage checkpoint will be compromised (Celeste, Fernandez-Capetillo, et al. 2003). Various proteins have been documented as being downstream effectors, and they typically possess 14-3-3 and BRCT domains which go on to recognise histones that have been phosphorylated (Taverna et al. 2007).

Phosphorylation of histones is not exclusively reserved for the DDR but can also play a role in transcriptional regulation. When H3S10, H3T11, and H3S28 are phosphorylated, they coincide with adjacent or proximal lysine acetylation, PTMs that are synonymous with gene expression. H3 acetylation in a GCN5 dependent manner has been shown to associate with the aforementioned phosphorylations (Lo et al. 2000). H3S28ph also plays a role in combination with H3K27ac in transcription activation by perturbing any nearby PRC reactions with chromatin, resulting in demethylation of H3K27, and subsequently, its acetylation. This would occur on adjacent residues on cis tails (Gehani et al. 2010). H3T11 and H3T6 have also been shown to displace methylation from H3K9 in a JMJD2C-dependent manner and encourage acetylation (Metzger et al. 2010). The same marks can also prevent H3K4 from being methylated, which would most likely result in transcriptional silencing in the local region. (Solier et al. 2009)

Several phosphorylation sites in the nucleosome have been connected to apoptosis. However, this section of the putative histone code remains unsolved as the exact combination required for inducing apoptosis has been contentious in recent years. H2AX on its most C-terminal residue, Tyrosine 142, can be phosphorylated, but this modification has been shown to decrease in the presence of DNA damage, is also putatively implicated in apoptosis (Xiao et al. 2008). Solier *et al* (2009) suggested three sites on two histones were crucial for the cell to undergo either repair of DNA, or to commit to apoptosis if the damage is too severe. The sites were H2AXS139, H2AXY142 and H2BS14; in the presence of phosphorylation at all three sites, the cell will be fated for apoptosis (Solier et al. 2009). Lastly, it has been documented that upon DNA damage of the nicking sort, apoptotic cells can undergo increased phosphorylation in a PKC δ -dependent manner at H3T45 (Hurd et al. 2009)

1.4.5. Histone Ubiquitination

Originally discovered by Goldstein et al. (Goldstein et al. 1975), ubiquitin has since been shown to be present in all eukaryotic cells, hence the name (it is ubiquitous). Unlike other PTMs, which involve a small adduct being added to a sidechain, ubiquitination involves the addition of one (or more) ubiquitin molecules, each with a molecular mass of 8.5 kDa (**Figure 1.5c**). A single ubiquitin can be added, or a chain of ubiquitin molecules. Ubiquitination can signal a protein for degradation but can also initiate or block interactions, change the location of a protein in a cell or alter the activity of a protein. Histones were the first proteins described to have been modified with a ubiquitin module (Goldknopf et al. 1975). Histone H2A and H2B are two of the most abundant proteins that carry ubiquitin sites: it is estimated that around 5-15% of H2A and 1-2% of H2B are ubiquitinated in the genome at any one time. There are three groups of catalysts, known

as ubiquitin ligases, that carry out ubiquitination on histone H2A: RNF168; RING1B (RNF2), which is a subunit of the PRC1 complex; and BRCA1/BARD1. Histone H2A has been found to be ubiquitinated in response to DNA damage. H2A K13/K15 is ubiquitinated by RNF8 and RNF168. ATM phosphorylates MDC1 which results in recruitment of RNF8, which then goes on to recruit RNF168 to the site of damage and the E2 ligase UBC13, leading to the synthesis of lysine63-linked ubiquitin (K63-Ub) chains at the site of damage (Yan 2012). Ubiquitination of H2A at different sites can also act as a gatekeeper to dictate the DNA repair mechanism. For example, BRCA1/BARD1 ubiquitination of H2A in the context of DNA damage promotes the usage of HR to repair DSBs (Densham et al. 2016). On the other hand, phosphorylation of H2AX will recruit RNF168 which will result in a docking site for 53BP1 to be constructed at H2AK13/K15 (Panier & Boulton, 2014). 53BP1 cannot dock at the site of a double-strand break if H2AK13ub/K15ub is not present. In addition to K13/K15, H2A has also been observed to be ubiquitinated at sites K118/K119 in response to damage. RNF20/RNF40 and PRC1 are the ubiquitin ligases responsible in this case (Pan et al. 2011). H2A K118/119 ubiquitination was previously implicated in having a role in repressing transcription (Yan 2012). Nonetheless, the recruitment of E3 ubiquitin ligase in γ H2AX dependent manner is necessary before HR-mediated or NHEJ-mediated repair of DNA damage can take place. Upon ATM phosphorylation of H2AX, MDC1 recognises the newly phosphorylated site and is also phosphorylated in a CK2-ATM-dependent manner. These ubiquitin chains allow for ubiquitin binding proteins to aggregate at the site of damage. One such protein is RAP80, which directly interacts with BRCA1, whose recruitment leads to HR-dependent repair. Alternatively, it is possible that 53BP1 will be recruited in a H4K20me2 dependent manner to the same sites, leading to HR-independent repair (Kolas et al. 2007).

PRC1-dependent H2AK119 ubiquitination has been found to reduce transcription at global levels. In addition to the DDR and transcription, H2A ubiquitination has been seen in the silenced X chromosomes of mammalian cells (Barr bodies)(Fang et al. 2004). H2B ubiquitination was demonstrated to play a role in chromatin demarcation wherein its absence resulted in other histone modifications pervading to other regions of the genome (Fierz, et al., 2011). Cells that did not possess the H2BK120 ubiquitination were shown to have defects in HR-dependent repair as it was crucial for the activation of the damage checkpoint (Nakamura, et al., 2011).

1.4.6. Lysine Propionylation

It is chemically feasible to modify a peptide in-vitro to aid with quantification. Derivatising a residue involves covalently attaching a functional group at an acceptor site. In this case, derivatising unmodified or monomethylated lysines with a propionyl group on the amine of the R-group. Derivatisation step pre-digestion has two main functions. The first is to ensure that there is uniformity and/or sufficient peptide length after digestion. For example, one can derivatise lysine residues so that regardless of whether there is a post-translational modification on a given lysine, the peptide sizes will remain identical to that of a peptide sequence that has not been modified. Derivatisation involves covalently modifying the sidechains of the amino acid so that they exhibit preferred properties e.g. derivatising an unmodified lysine residue so that it will not be digested with trypsin. This is because trypsin digests at lysine with no modifications or monomethyl, but not when there are larger adducts attached to the lysine side chain. Ideally, peptides should be longer than four amino acids otherwise they may be below the threshold of detection.

The second effect of derivatising the peptide with propionic anhydride is that it may have the effect of elongating the peptide, sequence permitting, but it will definitely make the peptide more hydrophobic, thus simplifying quantitative analysis of peptides and their PTMs (Sidoli et al. 2015) (**Figure 1.5i**). The analysis is simplified because the peptides will all be uniform in length after trypsinisation. The uniformity simplifies the quantification of a PTM. The frequency of the different PTMs signals can be tallied instead of calculated against a control peptide. Since the different signals corresponding to the different permutations are tallied, a direct percentage can be obtained. If there are peptides of varying sizes, the peptides will need to be normalised against a control peptide, which is commonly either a spiked extrinsic peptide or an endogenous peptide. For a full summary of known HPTMs, their sites and putative function, please refer to Table 1.1.

Table 1.1: Chart of Characterised Histone Residues Susceptible to Acetylation, Methylation, Phosphorylation and Ubiquitination and Histone Modifying Enzymes

Acetylation			
Histone	Site	Histone-modifying Enzymes	Proposed Function
H2A	Lys5	Tip60, p300/CBP	transcriptional activation
	Lys12	p300, ATF2	transcriptional activation
H2B	Lys12	p300/CBP, ATF2	transcriptional activation
	Lys15	p300/CBP, ATF2	transcriptional activation
H3	Lys20	p300	transcriptional activation
	Lys9	unknown	histone deposition
H3	Lys14	Gcn5, SRC-1	transcriptional activation
		unknown	histone deposition
		Gcn5, PCAF	transcriptional activation
		Esa1, Tip60	transcriptional activation
			DNA repair
		SRC-1	transcriptional activation
		Elp3	transcriptional activation (elongation)
		Hpa2	unknown
		hTFIIIC90	RNA polymerase III transcription
		TAF1	RNA polymerase II transcription
	Lys18	Sas2	euchromatin
		Sas3	transcriptional activation (elongation)
		p300	transcriptional activation
		Gcn5	transcriptional activation, DNA repair
		p300/CBP	DNA replication, transcriptional activation
		unknown	histone deposition
		Gcn5	transcriptional activation, DNA repair
		Sas3	transcriptional activation (elongation)
		p300/CBP	transcriptional activation
		p300/CBP	transcriptional activation
H4	Lys5	Gcn5	transcriptional activation
		DNA repair	
		Hat1	histone deposition
		Esa1, Tip60	transcriptional activation
			DNA repair
		ATF2	transcriptional activation
		Hpa2	unknown
		p300	transcriptional activation
		Gcn5, PCAF	transcriptional activation
		Esa1, Tip60	transcriptional activation
	Lys8		DNA repair
		ATF2	transcriptional activation
		Elp3	transcriptional activation (elongation)
		p300	transcriptional activation
		Hat1	histone deposition
			telomeric silencing
		Esa1, Tip60	transcriptional activation
			DNA repair
		Hpa2	unknown
		p300	transcriptional activation
H4	Lys16	Gcn15	transcriptional activation
		Esa1, Tip60	transcriptional activation
			DNA repair
		ATF2	transcriptional activation
		Sas2	euchromatin

Ubiquitination			
Histone	Site	Histone-modifying Enzymes	Proposed Function
H2A	Lys13		DNA Repair
	Lys 15		Dna Repair
	Lys119	Ring2	spermatogenesis
H2B	Lys120	UbcH6	meiosis

Methylation			
Histone	Site	Histone-modifying Enzymes	Proposed Function
H1	Lys26	Ezh2	transcriptional silencing
H2A	Arg3	PRMT1/6, PRMT5/7	transcriptional activation, transcriptional repression
	Arg2	PRMT5, PRMT6	transcriptional repression
H3	Arg8	PRMT5, PRMT2/6	transcriptional activation, transcriptional repression
	Arg17	CARM1	transcriptional activation
	Arg26	CARM1	transcriptional activation
	Arg42	CARM1	transcriptional activation
	Lys4	Set 7/9	transcriptional activation (tri-Me)
		MLL, ALL-1	transcriptional activation
	Lys9	Suv39h, Clr4	transcriptional silencing (tri-Me)
		G9a	transcriptional repression genomic imprinting
		SETDB1	transcriptional repression (tri-Me)
	Lys27	Ezh2	transcriptional silencing
			X inactivation (tri-Me)
		G9a	transcriptional silencing
	Lys36	Set2	transcriptional activation (elongation)
	Lys79	Dot1	euchromatin
H4			transcriptional activation (elongation)
			checkpoint response
	Arg3	PRMT1/6	transcriptional activation
		PRMT5/7	transcriptional repression
	Lys20	PR-Set7	transcriptional silencing (mono-Me)
		Suv4-20h	heterochromatin (tri-Me)
	Lys59	unknown	transcriptional silencing

Phosphorylation			
Histone	Site	Histone-modifying Enzymes	Proposed Function
H1	Ser27	unknown	transcriptional activation, chromatin decondensation
	Ser1	unknown	mitosis, chromatin assembly
H2A		MSK1	transcriptional repression
	Thr120	Bub1, VprBP	mitosis, transcriptional repression
H2AX	Ser139	ATR, ATM, DNA-PK	DNA repair
	Thr142	WSTF	apoptosis, DNA repair
H2B	Ser14	Mst1	apoptosis
		unknown	DNA repair
H3	Ser36	AMPK	transcriptional activation
	Ser10	Aurora-B kinase	mitosis, meiosis
		MSK1, MSK2	immediate-early gene activation
		IKK-α	transcriptional activation
		Snf1	transcriptional activation
	Ser28	Aurora-B kinase	mitosis
		MSK1, MSK2	immediate-early activation
	Thr3	Haspin/Gsg2	mitosis
	Thr6	PKCβI	
	Thr11	DIK/Zip	mitosis
H4	Tyr41	JAK2	transcriptional activation
	Tyr45	PKCδ	apoptosis
		unknown	mitosis, chromatin assembly
	CK2	CK2	DNA repair

1.5. Biological Mass Spectrometry

Since the turn of the century, mass spectrometry has become a useful tool for analysing biological molecules such as proteins and lipids. Proteomics is the large scale study of the proteins produced by living organism, groups of organism or system. With constant development of new techniques and refinement of protocols, proteomics is capable of elucidating chromatin dynamics and quantifying HPTMs in a multiplexed manner. Prior to this, most of what is known about HPTMs was learnt through immuno-assays and the scope of the information was limited. To further understand the advantageous of mass spectrometry-based methods in the area of HPTMs, it is necessary to understand the basic principles of mass spectrometry.

1.5.1. Introduction to mass spectrometry

The mass spectrometer is an analytical apparatus that allows for measurements of the isotopic composition of an element or the mass of molecules. It is a technology that can trace its origins to the late 19th century where the first mass spectrograph was generated. By the 1920s, what is considered to be the prototype for a modern mass spectrometer was developed by Arthur Dempster and Francis Aston. Prior to the mass spectrometer being able to analyse particles, the particles have to be both charged and in a gaseous state. The mass spectrometer measures the mass-to-charge ratio (m/z) of the ionised gas-phase particles that accelerate through a chamber containing an electric or magnetic field. The m/z can then be used to infer the elemental composition of the molecule. Since it is not the molecular weight that is measured but the mass to charge ratio of individual ions, it is highly improbable that the m/z value measured would be an integer due to the relative abundance of isotopes in nature.

Figure 1.6: Stages of Events within an LC-MS Orbitrap Mass Spectrometer

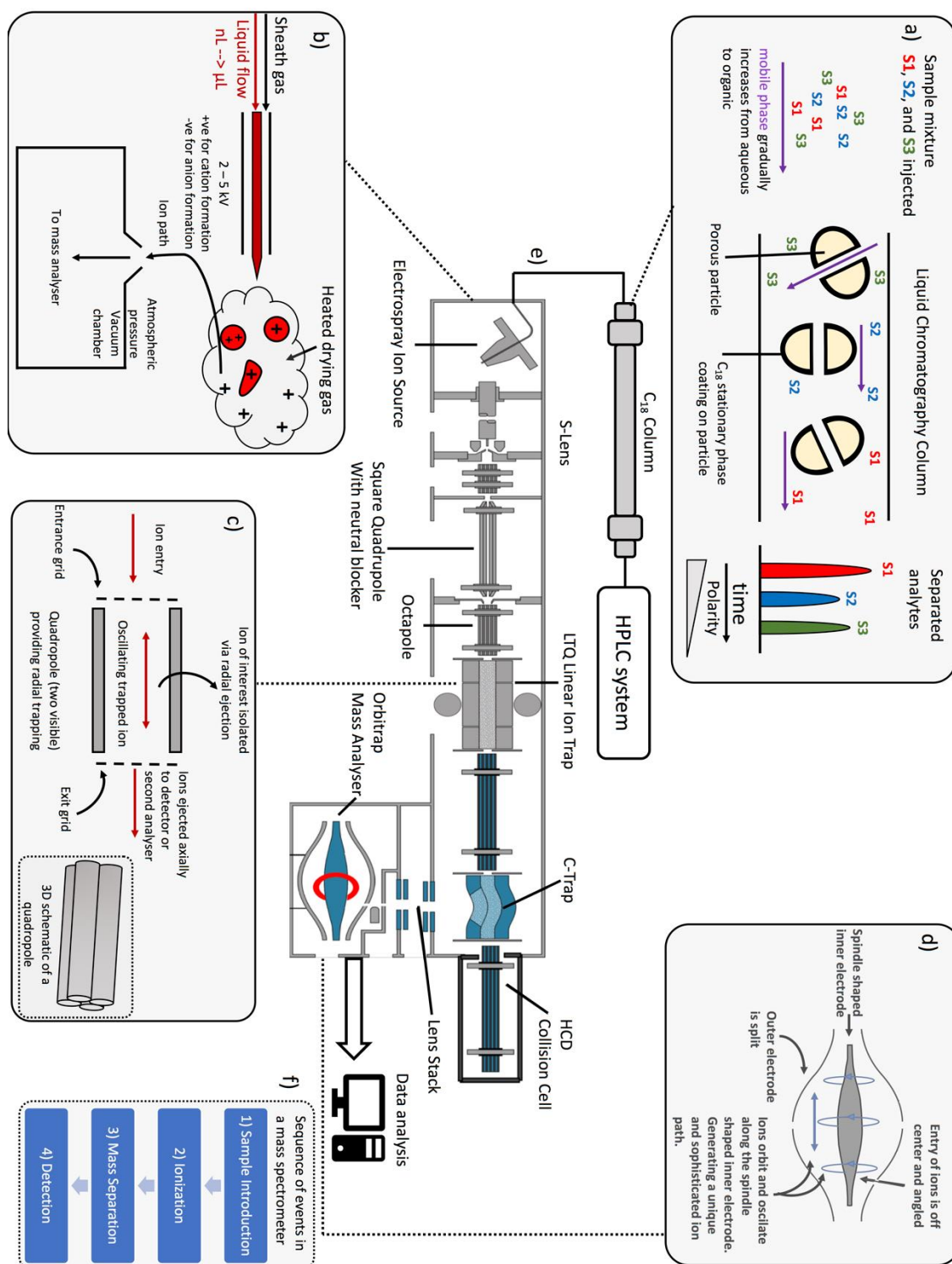


Figure 1.6 Legend: Stages of Events within an LC-MS Orbitrap Mass Spectrometer

LTQ Orbitrap Schematic Adapted from (Editors 2008). ESI and HPLC Adapted from (Greaves & Roboz 2013)

- a) Sample introduction: Analytes can be injected via direct infusion (not shown) or separating peptides based on their hydrophobicity. Hydrophilic peptides pass through the column the quickest due to the C₁₈ coating. Eventually, hydrophobic peptides pass through as the organic gradient gradually increases over time
- b) Ionisation: Many forms of hard and soft ionisation are available. Electrospray ionisation is suited for biological material with high polarity. Analyte in mobile phase is nebulised. Ions escape droplets and enter the mass analyser at an angle. At low flow rates, evaporation is enough for the gas to dry.
- c) Lateral view of quadrupole: Ions are trapped oscillating due to opposing electromagnetic rods encasing the ions. Ions can be ejected radially or axially. Linear ion traps have higher sensitivity compared to quadrupole ion traps as the ions do not cross paths as much.
- d) Orbitrap mass analyser: Spindle shaped electrode encased in an outer electrode. Ions enter at an angle for the purpose of creating momentum to orbit around the spindle while simultaneously oscillating laterally. The ion trajectory create unique paths which can be converted into spectra using Fourier Transformation
- e) Complete schematic of LTQ Orbitrap XL. From separation of analytes via HPLC to data analysis on an instrument attached computer
- f) Complete sequence of events of a mass spectrometer

Therefore, the m/z is almost always a non-integer and the number of significant figures given depends on the resolution of the instrument. The non-integer values highlights the precision to which the mass spectrometer can determine which element or molecule is present.

There are three components of a mass spectrometer. Firstly, an ion source that ionises molecules, transferring produce gas-phase ions. The second component of a mass spectrometer is the mass analyser, that sorts the ions based on their m/z ratio. Finally, a detector records the charge induced by an ion flying by or colliding with the surface (**Figure 1.6**) (Glish & Vachet 2003).

1.5.2. Ionization

To get to the mass analyser, ionization is a prerequisite for the analyte so that it is able to accelerate through an evacuated chamber filled with magnetic or electric fields. There are two types of ionisation: hard and soft ionisation. Hard ionisation leads to large quantities of residual energies, resulting in increased levels of fragmentation of the analyte. Therefore, hard ionisation is not suitable for macromolecules prepared from biological systems. It is ideal for inorganic compounds that have a lower molecular weight and can remain intact in a harsh environment. Soft ionisation results in relatively less fragmentation (Glish & Vachet 2003). Soft ionisation is the preferred method when dealing with macromolecules such as peptides and oligonucleotides. The two main ionisation techniques in proteomics are Electrospray Ionization (ESI) and Matrix-assisted laser desorption/ionisation (MALDI), which are both soft ionisation methods. They both cater to thermodynamically labile organic compounds so that they can be converted into the gas phase as a whole unit. The development of the aforementioned soft ionization

methods led to the 2002 Nobel Prize in Chemistry being awarded to John Fenn and Koichi Tanaka for devising ESI and LDI, respectively.

With ESI, the sample is introduced in solution. The solution flows through a capillary at atmospheric pressure and is then aerosolised when a high voltage is applied to the liquid sample, becoming a charged mixture of solvent and sample. Eventually, with enough energy, the droplet is desolvated (**Figure 1.6b**). Depending on the polarity of the voltage, the sample will either be protonated or deprotonated, resulting in a positively and negatively charged ion, respectively. As the size of the biomolecule increases, so does its susceptibility to receiving multiple protonations/deprotonation, which gives the ion a bigger charge and thus smaller m/z ratio. Multiple potential sites of de/protonation also result in different molecules of the same compound possessing different m/z measurements, which in the case of larger proteins, where isotopic resolution may not be achievable, can be deconvoluted to determine the charge-states. Higher charge states decrease the m/z ratio and increase the effective mass range of detection (Yamashita & Fenn 1983).

MALDI differs from ESI in that the sample is in the solid phase. There are three steps to MALDI: the analyte crystallises concurrently with a matrix that acts as a media to absorb high energy radiation from the irradiating laser, resulting in ablation and desorption of the sample from the matrix and, ultimately, conversion of the sample into ions in the gaseous state. Upon ionisation, the sample can then be accelerated through the magnetic/electric field. However, MALDI generates predominantly singly charged ions; this is in contrast to ESI, where the samples can be protonated and deprotonated at multiple sites (Karas & Hillenkamp 1988).

As mentioned, both ESI and MALDI are soft methods of ionising, making them suitable methods of ionisation for the analysis of biological material. They do differ in which applications they are suitable for. One example would be that ESI is better suited to analysing smaller molecules with lower m/z values, as the MALDI matrix results in low m/z background noise. For a more comprehensive comparison of the advantages and disadvantages of ESI and MALDI, and where they are both best deployed, refer to **(Figure 1.7)**.

Figure 1.7: Factors Determining Which Ionization Method To Use

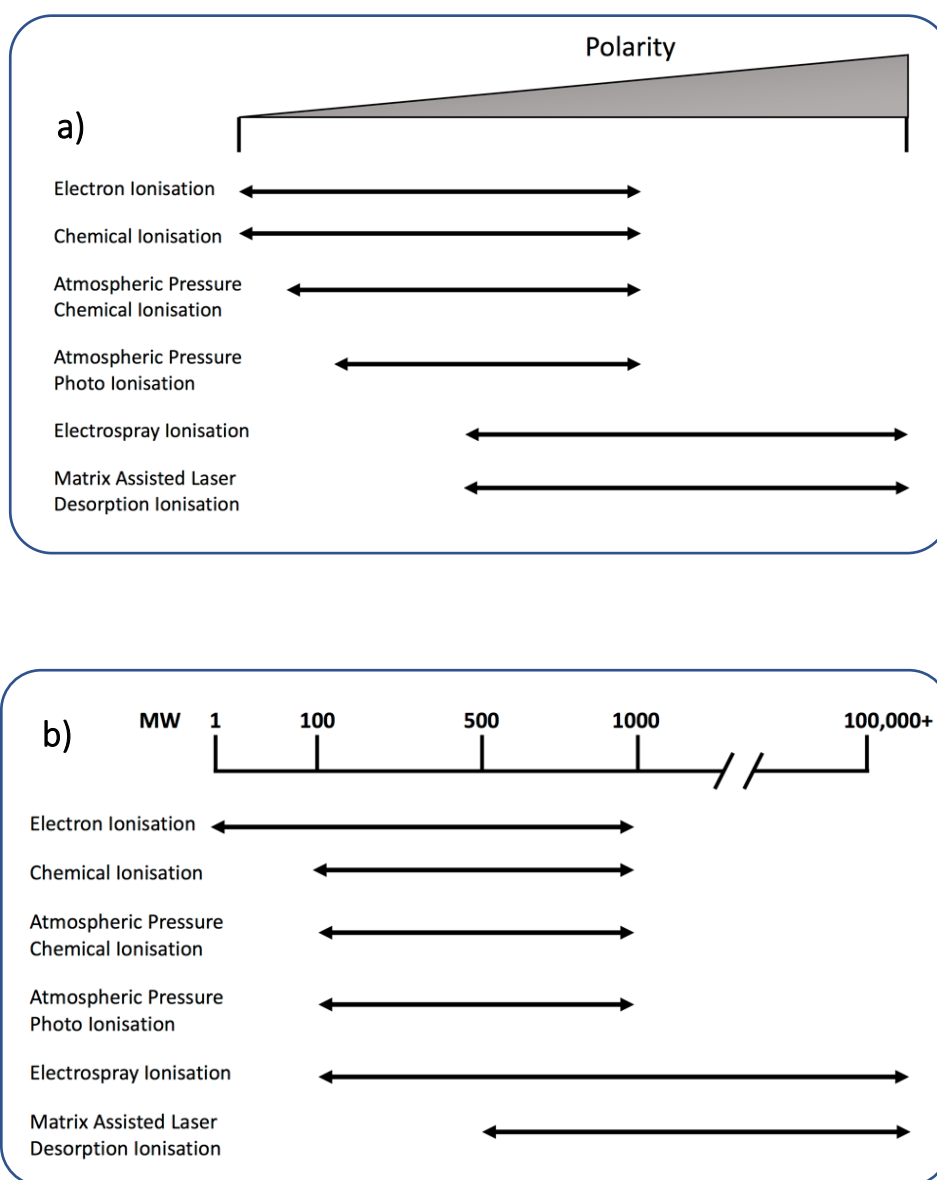


Figure 1.7 Legend: Factors Determining Which Ionization Method To Use (Greaves & Roboz 2013)

- a) **Polarity:** Depending on the polarity of the analyte in question there are a range of ionisation techniques that together can cover the full spectrum of polarity from organic small molecules to biomolecular polymers
- b) **Mass Range:** EI is able to ionise masses in the low double digit Daltons. CI, APCI, and APPI possess a similar 100-1000 Da range in ionizing capability. ESI has the highest range and in conjunction with being the softest ionisation method is therefore the most utilised in biological mass spectrometry. MALDI cannot ionise as well at the low range compared to ESI but it has great ionization efficiencies of >500 Da

APCI; Atmospheric Pressure Chemical Ionisation

APPI; Atmospheric Pressure Photo Ionisation

CI; Chemical Ionisation

EI; Electron Ionisation

ESI; Electrospray Ionisation

MALDI; Matrix Assisted Laser Desorption Ionisation

1.5.3. HPLC separation vs direct infusion

Depending on the source of the sample and its complexity, it might be necessary to separate a complex mixture with a High Performance Liquid Chromatography (HPLC) step before applying the sample to the mass spectrometer. Reverse stationary phase uses a column packed with hydrophobic nano-particles and depending on the properties of the peptides, they will either pass through the column quickly or slowly, with the more hydrophobic peptides being retained for longer. Essentially, a gradient to analyse the peptides has been generated so that they do not arrive at the ESI all at once (**Figure 1.6a**).

If a pure sample is being analysed, such as a sample obtained from peptide synthesis, then direct infusion without any further separation step would be a suitable method of analysis. The advantage would be that one could rapidly get confirmation of the mw of a substance. As the name suggests, direct infusion involves the loading of the analyte directly into the ESI through a capillary. However, for most other biological samples, it would be preferable to employ a miniaturised high-performance liquid chromatography (nanoLC). Standard-flow is an option. However, nano-flow enables one to work with much smaller concentrations of samples such as from immune-precipitation. In molecular and cell biology, seldom is a sample composed of only one type of molecule. Prior to some form of protein fractionation or enrichment, tens of thousands of different protein species will be present in the cell. After selection, for example via immunoprecipitation (IP), the vast majority of soluble proteins may have been excluded but non-specific artefacts will still be present in the sample. Moreover, the sample will have been digested with some form of protease (typically trypsin) during sample preparation, resulting in the generation of a variety of peptides with different biochemical properties. These peptides undergo further separation through the use of

HPLC (see above), ensuring that they do not all arrive at the ESI simultaneously. Reversed-phase chromatography is a system for separating out peptides based on their hydrophobicity and is employed in both nanoLC and HPLC systems. The stationary phase consists of alkyl chains covalently attached to beads packed within the column; due to the nonpolar nature of the alkyl chains, the stationary phase is hydrophobic. On the other hand, the mobile phase is an aqueous solution. Hydrophobic peptides will be attracted to the environment created by the alkyl groups and will adsorb to the stationary phase, whereas hydrophilic peptides will remain in their preferred environment i.e., the aqueous mobile phase (**Figure 1.8**). As a consequence, hydrophilic peptides exit the column first in the form of an unbound flow-through fraction. The hydrophobic peptides are eluted from the column by applying a concentration gradient of organic solvent, i.e., non-polar solution. The least hydrophobic peptides, which adsorb most weakly to the alkyl groups, will be eluted at lower concentrations of organic solvent. Higher concentrations of organic solvent will be required to elute the more hydrophobic peptides from the column.

In proteomics, separation of trypsinised peptides is crucial to generate high signal to noise data, which is essential for generating results that will have quantitative relevance. The nanoLC connected to the ESI method provides a stable flow rate, ensuring that the droplet size is consistent and minimising any adverse effects on the ionization efficiency. For targeted MS, where m/z correlating to peptides of interest are picked (**Figure 1.8**), to ensure the highest degree of accuracy and precision, consistent retention times of peptides is a necessity. A modern and sophisticated nanoLC can provide this. Otherwise,

some scans of some targeted masses will be missed if they drift out of the window, thereby affecting results downstream (Rudnick et al. 2010) (Mason et al. 2005).

Figure 1.8: Targeted LC-MS-MS

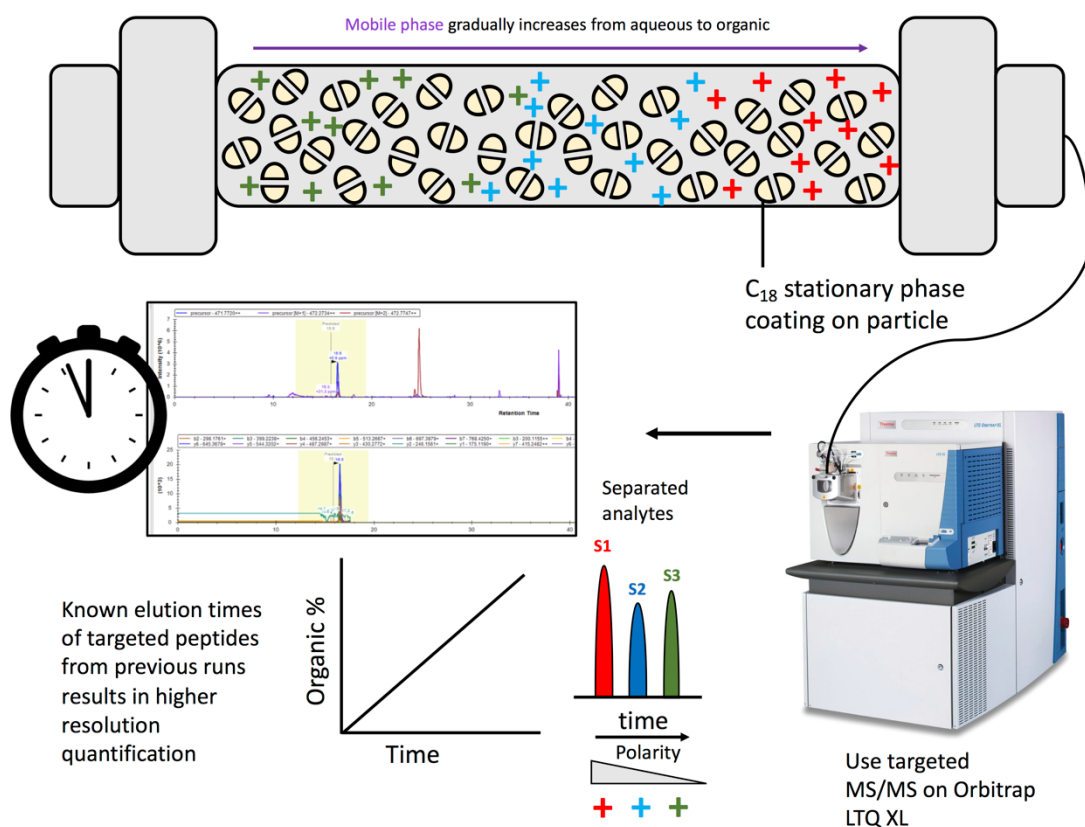


Figure 1.8 Legend: Targeted LC-MS-MS: Data Dependent Acquisition

Targeted LC-MS-MS can increase the levels of proteomic quantification. Samples are initially acquired a data independent acquisition (DIA) method. The prior run can be used to identify the presence of analytes of interest and when elution times for the analytes from the C₁₈ column.

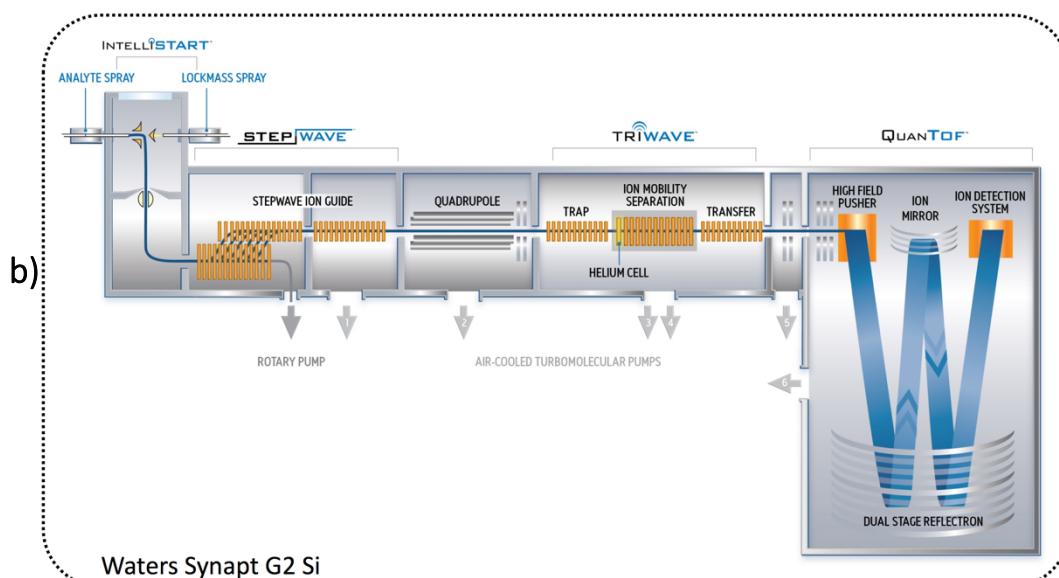
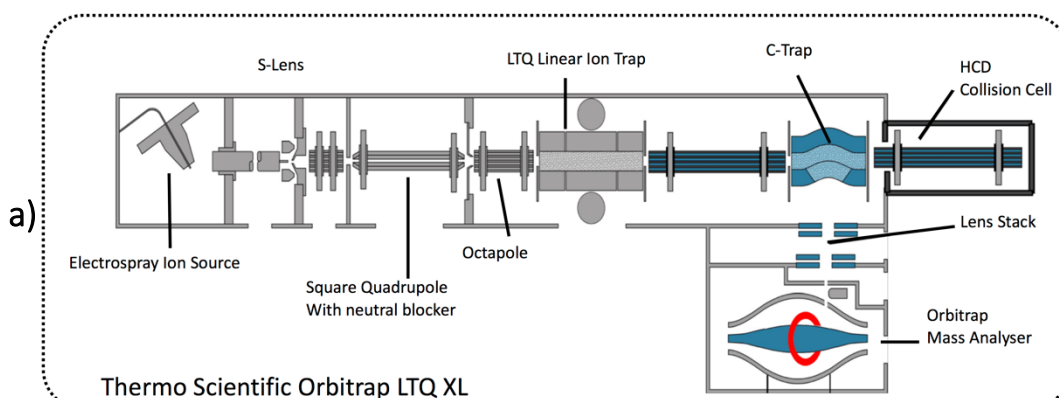
The second run involves a Data Dependent Acquisition i.e. instructing the instrument with which masses to target and when. This method enables high resolution quantification which can be used to compare expression levels of a peptide or a PTM of a residue of interest.

1.5.4. Orbitrap

The Orbitrap-XL utilized in this work is a hybrid instrument, incorporating a linear ion trap and an Orbitrap. The Orbitrap is composed of an inner electrode that is spindle-like in shape within a cylindrical electrode. Due to the geometry of the charge from the two electrodes, ions remain trapped orbiting the spindle-shaped electrode. The ions also oscillate around the spindle pole to pole. **(Figure 1.9a)**. The oscillation generates an image current on the detector plate which the machine records. Using an algorithm derived from Fourier transformations of the recorded image, the m/z ratio can be deduced by the frequency of the image currents and a mass spectrum can be generated (Makarov 2000).

A quadrupole mass analyser is composed of four rods orientated in a parallel conformation. The quadrupole can use the m/z ratio of sample ions to sort them based on their flight path whilst oscillating around the electric fields applied to the rods. A linear ion trap will trap radially by utilising a quadrupole configuration. The second set of electrodes is then used to confine the ions axially. Linear ion traps are useful for confining the ions into a "well" which can then be used to filter the ions based on their m/z . This allows the machine to store a larger amount of ions and perform faster scans but with reduced resolution compared to the orbitrap (Soldi et al. 2014). An Orbitrap instrument possesses a high mass accuracy, is has high resolving power and can process samples over a large m/z ratio range. However, it does not provide the highest sensitivity or dynamic range compared to other mass analysers **(Figure 1.9c)** (L. Sun et al. 2012).

Figure 1.9: MS/MS Compatible Instruments: Orbitrap vs Q-IMS-TOF



c)

	Synapt G2 Si	Orbitrap LTQ XL
Resolution	Sensitivity mode: 10,000 Resolution mode: 20,000 High Resolution mode: 40,000	FWHM: 60,000 @ m/z 400 Min: 7500 Max: 100,000
Acquisition rate	30 scans per second	1 scan at 60,000 resolution @ m/z 400 per sec
Mass Range	Resolution mode: 20 - 100,000 High Resolution mode: 20- 32,000	m/z 50–2000; m/z 200–4000
Mass Accuracy	> 1ppm RMS with internal lock mass between 150 – 900 m/z ;	< 3 ppm RMS external calibration; <2 ppm RMS internal calibration
Dynamic Range	High resolution mode: 4 orders of magnitude	>10,000 between mass spectra >4000 between highest and lowest detectable in one spectrum

Figure 1.9 Legend: MS/MS Compatible Instruments: Orbitrap vs Q-IMS-TOF

Adapted from (Editors 2008) and (Waters Corporation 2016)

- a) Thermo Scientific Orbitrap LTQ XL: Orbitraps have the second highest resolution of mass spectrometers. Second only to Fourier Transform Ion Cyclotron Resonance. Both instruments use frequency and Fourier Transformation to give a higher mass accuracy than a Quadrupole-Time of Flight Instrument. The Orbitrap also has high mass resolution compared to Q-TOF based system.
- b) Waters Synapt G2 Si: While the resolution of this Q-TOF instrument is not as high as that of the Orbitrap, the Synapt G2 Si has an extra layer of resolving: the Ion Mobility Separator that differentiates ions not based on charge but on shape. Essentially three dimensions of resolution. It also trumps the Orbitrap in regards acquisition rates. The Synapt G2 Si also utilises a “W-Wave” in the TOF cell to increase the mass accuracy compared to a linear TOF cell.
- c) Table of Comparison of Manufacturers’ Specifications

1.5.5. Tandem Mass Spectrometry: pseudo-Selected Reaction Monitoring

Even though mass spectrometry can in some cases obtain the elemental composition of a molecule from the m/z ratio, it is not possible to obtain the amino-acid sequence from a peptide without fragmentation. To obtain the amino-acid sequence of a particular peptide, it is necessary to isolate and fragment the corresponding peptide ion (precursor ion) with the aid of the linear ion trap. With a combination of different fragment ions being generated from a given precursor ion, one can infer the amino-acid sequence. This technique is known as tandem mass spectrometry and it allows for precise identification of peptides and localisation of PTMs to any given amino-acid residue (Glish & Goeringer 1984) (Steen & Mann 2004). Using tandem MS allows for high levels of specificity to be attained. Two mass selectors are utilised by setting them to transmit targeted pairs of parent and fragment ions. A triple-quadrupole instrument enables increased sensitivity with the use of fragmentation. Two mass analysers are utilised; the first to gate the precursor ion and the second to gate fragment ions. In conjunction with high frequency cycles available on Q-TOFs, the two filtering steps lead to ultrahigh sensitive quantitative analysis (**Figure 1.10**). The precursor and fragment are together known as a 'transition'. Tandem mass spectrometry has become a necessity for quantitative proteomics.

Figure 1.10: Tandem Mass Spectrometry (MS/MS) and Applications

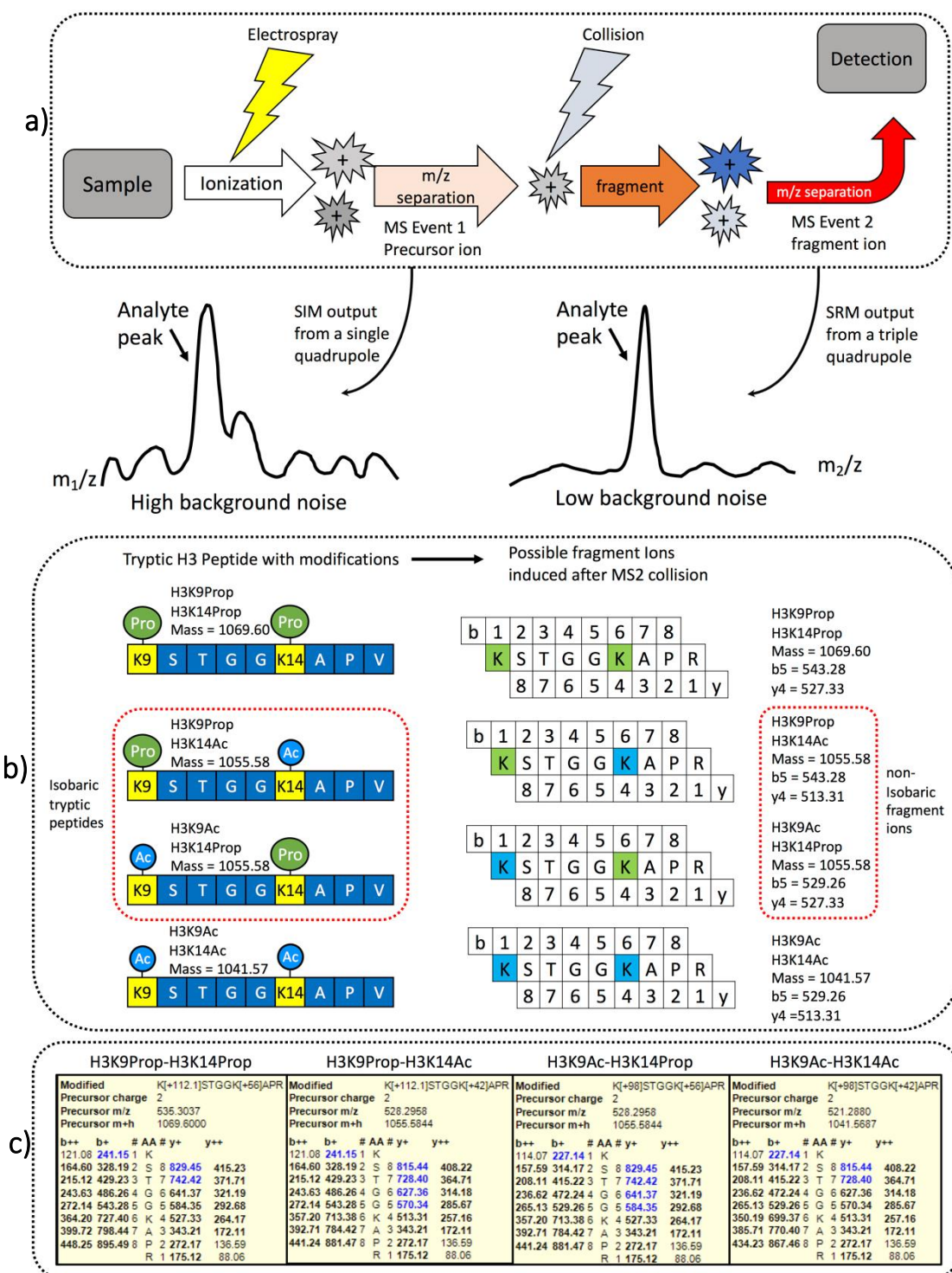


Figure 1.10 Legend: Tandem Mass Spectrometry (MS/MS) and Applications

- a) (i) In tandem mass spectrometry, precursor ions are initially selected based on their m/z . The ions are fragmented in high energy collisions. The product ions can be further selected. (ii) There are high levels of specificity with increased sensitivity compared to Single Ion Monitoring alone. (i) Tandem mass spectrometry can also help distinguish overlapping peaks in a spectrum or reduce chemical background noise.
- b) Resolving Power of MS/MS. A tryptic histone H3 peptide with two lysine residues has four permutations when it can only be modified by propionylation and acetylation. Therefore two of the possible peptides possibilities are isobaric i.e. carrying one of each modification on either peptide (red dotted square). Without tandem mass spectrometry the isobaric peptides could not be distinguished. However, with MS/MS, CID results in the generation of *b* and *y* ions, where residues break off at the C and N terminal respectively. Product ions carrying only one of the modifications will help to distinguish which modification is carried.
- c) Table of Masses For All Product Ions of Tryptic H3 Peptide

In canonical selected reaction monitoring (SRM), the selected transitions are gated by quadrupoles, allowing 100% duty cycle for a given m/z . However, the Orbitrap is not capable of SRM, lacking quadrupoles for precursor and fragment ion selection, therefore the equivalent targeted experiment is referred to as pseudo-SRM (pSRM). The reduced duty cycle makes the Orbitrap pSRM less sensitive than SRM. In contrast to SRM on a triple quadrupole (QQQ), pSRM benefits from high resolution detection of precursors and/or fragment ions, helping to detect close to isobaric interference (Gallien et al. 2013)

1.5.6. LC-MS/MS compared to Immunoassays

Immunoassays have been well established for several decades (Yalow & Berson 2007).

Rosalyn Yalow pioneered the use of antibodies and received the 1977 Nobel Prize in Physiology or Medicine for her work on developing radioimmunoassays. Immunoassays have been the standard for assessing protein expression and PTMs in biomedical/biochemical research for several reasons. Immunoassays give simple readouts of clinical assays or if a protein silencing has been successful. The antigen, which is the protein of interest, possesses epitopes that a specific antibody will recognise. The antibody specific for a protein of interest will have been raised in a known species, e.g. mouse, so a secondary antibody from a different species could be used, e.g. rabbit, that recognises IgG from the host species. The secondary antibody will be tagged with some form of reporter molecule. A common reporter tag is horseradish peroxidase (HRP), which fluoresces when it comes into contact with its substrate in the presence of an enhanced chemiluminescence (ECL) solution. Such a system can be utilised in western blotting, immunofluorescence, ChIP-on-chip, ChIP-seq and ChIP-qPCR. Immunoassays prove themselves useful in several ways: they are widely used around the scientific and medical community so libraries of extensive protocols and techniques are available. Also,

immunoassays do not have a steep learning curve so it is easier to train an individual on the rudimentary and common steps. They also provide decent sensitivity for analysis: modern microscopy techniques are able to detect single protein molecules (Etheridge et al. 2014). Antibodies are also able to cope with complex protein configurations e.g. a tried and tested antibody will be able to pull down a single species of protein from a clarified cell lysate. Financially, immunoassays have a low barrier of entry for a typical lab group. However, as proteomics has been improving its reputation as a formidable tool for analysing protein expression, immunoassays have had theirs tarnished in recent years due to antibody non-specificity. When raising antibodies, there will be inconsistencies from even the same individual animal if it is bled on different days. Batch-to-batch variation of the same antibody from the same manufacturer is a described issue. Moreover, a lack of rigorous antibody verification is a common problem with commercial antibodies. Traditionally, antibodies have been used to assess histone PTMs. However, in recent years, this approach has come under increasing scrutiny: numerous epigenetic antibodies that were used in immunoassays in previously published studies were found to be non-specific in their binding, which can obviously result in false positives (Egelhofer et al. 2010). There is no industry standard across manufacturers for the same protein of interest and so there is no consistent result for the same experiment across different labs (Tjeertes et al. 2009). In contrast, once a standard technique is established in LC-MS/MS, it is more reproducible as there is substantially less variability between a mass spectrometer of the same model and make than there is for antibodies (Paulovich et al. 2010). Some proteins, for example H2AX or 53BP1, are well characterised and therefore supply and demand states that there will be numerous antibodies available for that protein so there will be less chance of error. However, for the understudied proteins of

interest, the antibody availability would be scarce. The scarcity will require raising an antibody for the antigen which has greater potential for non-specificity. The non-specificity can result in misleading results downstream of the immunoassay. LC-MS/MS has no such inherent issues. Even though the use of mass spectrometry in molecular biology has become increasingly popular in recent years, compared to immunoassays, there are not as many established protocols. However, there is an increasing library of mass spectrometry-based protocols. The advantage is that the proteomic protocols are more reproducible across lab groups compared to immunoassay protocols (Tabb et al. 2010) (Eyers & Gaskell 2014) (Sechi 2007)

Not only do antibodies have specificity issues, they are also limited in the questions that they can answer. With mass spectrometry, it is possible to quantify protein isoforms amongst other isoforms which are not possible with immuno-based methods. In addition, mass spectrometry gives a high-throughput of HPTM quantification. With antibody-based systems, one would need to know which particular proteins and modifications they would be interested in prior to starting the immunoassay as the relevant antibodies would need to be purchased beforehand. However, mass spectrometry detection is not limited to what is available in the lab stock or manufacturers catalogue. LC-MS/MS allows for data-independent methods so that a whole data stream of what is and is not present in the sample can be generated. Due to the high throughput analysis that mass spectrometry affords, which particular HPTMs change in response to damage on a temporal scale can be determined. With the addition of SILAC (Stable isotope labeling by amino acids in cell culture) based systems, whether the modification changes are new modifications or old modifications are determinable. Which histone variants are present at the site of damage

can be checked simultaneously, rather than one or two (depending on a number of channels available with microscopy). We can see what the temporal chromatin dynamics are at the site of damage, which yields answers to questions such as whether more repressed chromatin exists later in the DNA repair process (Goodarzi et al. 2008). If an increase in repressive PTMs are found at later time points, we can then ask if those PTMs are old modifications or new modifications with heavy labelled methionine and arginine. As outlined, mass spectrometry is a high resolution, unbiased technique that can generate large amounts of data rapidly. It makes sense that the next step in epigenetics research would be to utilise state of the art proteomic techniques. A comparison of immune-assay against LC-MS is laid out in **Table 1.2**.

Table 1.2: Comparison of the capabilities and feasibility immuno-assays and LC-MS

Immuno-assay		Liquid Chromatography-Mass Spectrometry
Cost	Depending on the assay, relative cheap to set up e.g. western blotting. However can also be expensive to set e.g. cost of a sophisticated fluorescent microscope	Relatively expensive to set up. An instrument from any manufacturer can be prohibitively expensive for most groups. Reagents, maintenance and space required can be expensive
Ease-of-use	Very easy to train new members on most techniques	Steep learning curve compared to most immuno-assay techniques
Live Cell Work	Possible	Not possible
Design of Experiment	Bias: requires prior knowledge and hypotheses to utilise a given antibody	Non-bias: can discover new interacting partners or PTMs with data-independent acquisition method
Feasibility of scaling up experiment	Does not scale well for most assays and costs increase substantially	With the use of HPLC, tens if not hundreds of peptides can be targeted and analysed directly
Differentiate PTMs spatially	Can be an issue if raised antibody recognises modified residues proximal to epitope site	Use of b and y ions can determine which specific amino acid residues carry a PTM
Differentiate PTMs temporally	Pulse chase experiment can determine loss or gain of PTM but not determine new or old	SILAC can determine whether proteins or PTM are new or old
Less-characterised proteins	Pool of commercial antibodies may be smaller. May even require raising antibodies. Potential for trial and error and less reproducibility between batches	As long as peptides can ionize efficiently, there is no difference results in regards to well studied or less studied peptides.

1.6. HPTMs in the DDR; An LC-MS/MS Challenge (preface to results)

To summarise this introduction, three components of biochemistry have been explored in detail, with the aim of marrying them together: DNA damage response, histone post-translational modifications and mass spectrometry. Understanding the DDR can eventually pave the way for leading to better treatments for cancer, infertility and neurodegenerative diseases. The Histone Code, as coined by David Allis, has given the genome a whole new dimension of complexity when comparing to the transcriptome alone (Strahl & Allis 2000). It plays a role in a long list of biological functions, including the DNA damage response. However, with the different permutations that are possible with the histone code, new systems are required to answer more complicated questions. The aim of this thesis is to develop a mass spectrometry method that can be utilised to answer questions of the DDR HPTM changes that can occur on a spatiotemporal scale. Histones that have been prepared from the nucleosome indiscriminately can answer questions at a global level. However, if we were to select for histones specifically from damaged chromatin, then the HPTM changes identified would reflect what is taking place at a local level. Several protein tagging options are explored to develop such a technique, which will be outlined in later chapters. In conjunction with a pulse-chase experiment, questions can be asked on a time scale of when a particular HPTM is introduced and whether its presence overlaps or does not coincide with another protein or HPTM. Determining what is present or absent alongside a given HPTM would help to identify potential interaction partners, thus providing mechanistic insights into the relationship between HPTMs and DDR proteins. Here, we aim to analyse as many HPTMs that were reported in previous studies to change as a result of the DDR (**Table 1.1**) and ask whether we would see similar results with the developed proteomics method.

Chapter 2: Methods and Materials

2.1. Culturing of Cells and cell lines used

Mouse Embryonic Fibroblasts (MEF), Human U2-OS and HeLa cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% foetal calf serum, 2 mM L-glutamine, 100 U/mL penicillin, and 100 µg/mL streptomycin. Epstein-Barr virus-transformed wild-type (i.e. from an unaffected sibling of a patient; generous gift from the O'Driscoll lab) lymphoblastoid cells (LCLs) were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 15% foetal calf serum, 2mM L-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin.

Table of cell lines used in experiments

Cell Line	Origin	Immortalised
LCL Wild Type	Lymphoblastoid	EBV
U2-OS	Bone Osteosarcoma	Cancer
53BP1:3xFLAG U2-OS	Bone Osteosarcoma	Cancer
HeLa	Cervical Cancer	Cancer
MEF Wild Type	Mouse Embryo	TERT
MEF H2AX ^{-/-}	Mouse Embryo	TERT

2.2. DNA Damage: Ionising Radiation

Irradiating U2-OS Cells: DMEM media was aspirated and replaced with 5 ml Trypsin (0.05% EDTA) (GIBCO#25300) and the cells were incubated at 37 °C until they detached from the flask. Cells were spun down at 1500 rpm for five minutes and resuspended in 50 mL of DMEM media in a 50 mL conical centrifuge tube. U2-OS cells were irradiated with a dosage of 20 Gy using a ¹³⁷Cs source. Cells were immediately placed in a 37°C incubator and were left to recover for either 5 minutes, 30 minutes, or 60 minutes. Upon

recovery, cells were spun down and were immediately flash frozen in liquid nitrogen and stored at -80°C before sample preparation.

Irradiating LCLs: Cells from different flasks were collated together into suspension. Cells were then spun down at 1500 rpm for five minutes and suspended in new RPMI in a 15 mL conical centrifuge tube. LCLs were irradiated with a dosage of 20 Gy using a ^{137}Cs source. Cells were immediately placed in a 37°C incubator and were left to recover for either 5 minutes, 15 minutes, 30 minutes, 60 minutes or 24 hours. Upon recovery, cells were spun down and were immediately flash frozen and stored at -80°C before sample preparation.

2.3. Nuclear Extract

Modified from Lin et al. (Lin et al, 2012). Nuclei Isolation Buffer 250 (NIB-250) was prepared: 15 mM Tris-HCl (pH 7.5), 60 mM KCl, 15 mM NaCl, 5 mM MgCl_2 , 1 mM CaCl_2 , and 250 mM Sucrose. NIB-250 was supplemented with the following phosphatase and protease inhibitors:

For acid extractions in chapter 3:

- 10 mM Na-butyrate
- 0.5 mM AEBSF (4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride)
- 2 μM leupeptin
- 1 μM Pepstatin A
- 1 μM aprotinin
- 2 mM PMSF (phenylmethane sulfonyl fluoride)
- 5 mM NaF
- 5 mM NEM (N-Ethylmaleimide)
- 2 mM Na-orthovanadate

- 60 mM β -glycerophosphate,

For Micrococcal nuclease digestions preceding immunoprecipitations in Chapter 4:

- 10 mM Na-butyrate
- 0.5 mM AEBSF
- 2 μ M leupeptin
- 1 μ M Pepstatin A
- 5 mM NaF
- 5 mM NEM

NIB-250 with 0.3% NP-40 was added to the cells at a ratio of 10:1, by volume of cell pellet. Cells were incubated on ice for 5 min and then subsequently spun down at 600 rcf for five min at 4°C. The supernatant was discarded and then NIB-250 without NP-40 was added to the cell pellet and spun down again, twice. The remaining pellet consisted of isolated nuclei. The isolated nuclei were vortexed slowly and 0.4N sulphuric acid was added to the cells at a final ratio of 5:1. The nuclei-H₂SO₄ suspension was left to incubate on ice for 60 minutes and then spun down at the centrifuge's maximum speed for 5 minutes at 4°C. The supernatant was transferred to a microcentrifuge tube. To the supernatant, 100% trichloroacetic acid (TCA) was added to give a final concentration of 20% TCA. The mixture was left to precipitate overnight at 4°C. The precipitate was then spun down at the microcentrifuge's maximum speed. The pellet was resuspended in chilled acetone with 0.1% hydrochloric acid. The sample was then spun down, washed twice with acetone, left to air dry, and finally dissolved in water.

2.4. Acid Extraction of Histones

Modified from Lin et al. (Lin et al, 2012). The isolated nuclei were vortexed slowly and 0.4N sulphuric acid was added to the cells at a final ratio of 5:1. The nuclei-H₂SO₄ suspension was left to incubate on ice for 60 minutes and then spun down at the centrifuge's maximum speed (5,000 rpm; Eppendorf Centrifuge 5810) for 5 minutes at 4°C. The supernatant was transferred to a microcentrifuge tube. To the supernatant, 100% trichloroacetic acid (TCA) was added to give a final concentration of 20% TCA. The mixture was left to precipitate overnight at 4°C. The precipitate was then spun down at the microcentrifuge's maximum speed (13,500 rpm; Eppendorf Microcentrifuge 5424). The pellet was resuspended in chilled acetone with 0.1% hydrochloric acid. The sample was then spun down, washed twice with acetone, left to air dry, and finally dissolved in water.

2.5. Micrococcal nuclease treatment

Upon completion of nuclei isolation, the DNA concentration was measured: 10 µL of nuclei isolation was diluted in 89 µL ddH₂O, 1 µL 10% SDS was also added to give a final concentration of 0.1% SDS. The sonicator (diagenode Bioruptor® Pico) was switched on prior to use, to allow cooling of the device. The sample was sonicated for 10-15 cycles 30 seconds on, with 30 seconds off between each sonication. The number of cycles depended on the viscosity of the isolated nuclei dissolved in 0.1% SDS solution. A NanoDrop Microvolume spectrophotometer was used to measure the ratio of absorbance at 260 nm and 280 nm (A₂₆₀/A₂₈₀) to determine the DNA concentration of the sonicated isolated nuclei.

From the original unsonicated mixture, 500 – 2000 µg of chromatin was aliquoted into a new tube. The sample was spun down at 400 rcf and resuspended in digestion buffer:

- 0.25 M sucrose
- 50 mM Tris-HCl pH 7.9
- 5 mM CaCl₂

The following inhibitors were also added:

- 1 mM NaF
- 5 mM Na-Butyrate
- 20 µM NEM
- 0.2 mM AEBSF
- 10 µM Leupeptin
- 1 µM Pepstatin A.

Freeze dried micrococcal nuclease (Mnase) was resuspended in 0.25 M sucrose and 50 mM Tris-HCl. 1 U (Worthington Unit) was used to digest 150 µg of chromatin-bound DNA into mono- and dinucleosomes. The samples were placed in a micro-shaker-dry bath set at 37 °C for 50-60 minutes at 1500 rpm. EDTA was added to each reaction for a final concentration to each sample to quench the Mnase digestion. The samples were spun down (11,000g x 10 min) and the supernatant was collected (S1).

To release nucleosomes from more compact regions of the genome, the pellet was resuspended in 500 µL lysis buffer:

- 0.5 M NaCl
- 1 mM Tris-HCl pH 7.4
- 0.2 mM AEBSF
- 5 mM Na-Butyrate
- 20 µM NEM

- 1 mM NaF.

The samples were then shaken vigorously on a vortex for 30 minutes. The samples were then spun down at the tabletop microcentrifuge's maximum speed, ~ 14,000 RPM. The supernatant was then collected: S2.

The DNA concentration of both supernatants was measured using the NanoDrop, measuring the A260/280: the supernatants were diluted into 1% SDS and 1X New England Biolabs loading buffer, shaken vigorously for 15 minutes and then heated in a dry bath at 85°C for 15 minutes.

2.5.1. Agarose gel electrophoresis

To check the efficiency of the Mnase digestion, approximately 1.5 µg of DNA was loaded onto a 1% agarose gel with no ethidium bromide in 0.5 M TAE (Tris base, acetic acid and EDTA). The gel was run at 50 Volts for 45 minutes, or longer if required for better separation. The gel was then shaken in 1 in 20,000 Ethidium Bromide solution. DNA released from S1 should be in the size range of 200-400 base pair: denoting mainly mono- and dinucleosomes. S2 will have a larger range of DNA size as the Mnase would not have digested the DNA completely. A Mupid mini-gel electrophoretic unit was used to run the gels.

2.6. Immunoprecipitation of γH2AX

Upon digesting chromatin with Mnase, the supernatants S1 and S2 were pooled together for a ratio corresponding to approximately 1:2 respectively in protein content. The additional buffer was added to S1 and S2 mixture to ensure total salt concentration remained around 200 mM. An antibody concentration of 0.25 µg of γH2AX antibody (Abcam ab81299) was incubated overnight at 4°C with for every 1 mg of nucleosomes. A

total of 1.5 mg of protein was used to pull down γ H2AX. 10-20% percent of the starting material was retained as input. The next day, 10 μ l magnetic IgG beads (Dynabeads® Protein G for Immunoprecipitation) were washed twice in PBS, and once in incubation buffer. The beads were then left to incubate with the Mnase-digestion and γ H2AX antibody for 60 minutes. After incubation, a magnetic stand was used to collect the flow through which was to be kept for analysis with western blot, to determine depletion of γ H2AX. The beads were washed for five minutes for a total of five minutes in wash buffer:

- 500 mM NaCl
- 50 mM Tris
- 1% NP40
- 1% Tween

The beads were agitated fifteen minutes in Laemmli buffer. The beads were then boiled beads in the Laemmli buffer. The Laemmli buffer supernatant was collected using the magnetic stand. 2-Mercaptoethanol (β ME) was then added to the Laemmli buffer to give a final concentration of 5% β ME. The beads were boiled once more and then loaded onto an SDS-PAGE gel.

2.7. SDS-PAGE and Western Blot

BIO-RAD Mini-PROTEAN® Tetra Vertical Electrophoresis Cell was used for running SDS-PAGE electrophoresis gels and for transferring proteins onto membranes. 15% polyacrylamide gels were used to run and separate proteins:

Running Gel (1x)	Stacking Gel
1.8 ml ddH ₂ O	2.5 ml ddH ₂ O
4 ml 30% Bis-Acrylamide	1 ml 30% Bis-Acrylamide
2 ml 1.5M Tris pH 8.8	1.25 ml 0.5M Tris pH 6.8
80 µl 10% SDS	50 µl 10% SDS
80 µl 10% APS (Ammonium persulphate)	50 µl 10% APS
8 µl TEMED (Tetramethylethylenediamine)	4 µl TEMED

Gels were run at fixed voltage, 200 V for 60 minutes. Once the proteins were separated, the proteins were then transferred onto nitrocellulose membrane (0.45 µm, Bio-Rad, Catalogue number: 1620115) at fixed current 400 mA for 60 minutes. The membranes were subsequently stained with Ponceau S (Sigma-Aldrich) to check for complete transfer of proteins.

After imaging the Ponceau S stained membranes, the membranes were incubated overnight on a roller for agitation at 4°C in 0.2% Tween-TBS (TBST) with 5% milk powder for blocking. The next day, the TBST-milk solution was discarded and replaced with TBST-milk solution containing the primary antibody for probing a protein of interest. The primary antibody solution was left to incubate for an hour at room temperature. The membranes were then washed for 5 minutes in TBST, three times. Afterwards, the membranes were incubated with a secondary antibody that is specific for the species that the primary antibody was raised in. Secondary antibody incubation was at room temperature for 60 minutes. This incubation was followed by the second round of three 5 minutes TBST washes. Pierce™ ECL Western Blotting Substrate was then used for detection via chemiluminescence reaction with the horseradish peroxidase-conjugated

to the secondary antibody. Membranes were exposed onto x-ray film (GE Healthcare Amersham Hyperfilm ECL) for a short period of time ranging from 10 seconds to 5 minutes. The x-ray film was then developed using a Photon Ecomax Automatic X-ray Developer

2.8. Derivatization and digestion of peptides

2.8.1. In-solution derivatisation

Modified from Garcia et al. (Garcia et al. 2007). 5 µg of chromatin obtained from either acid extraction or Mnase digestion was resuspended in a volume of 9 µL, 1 µL of 1M tetraethylammonium bromide (TEAB) was added to the sample to buffer the subsequent reaction. 1 µL of 1% propionic anhydride (in acetonitrile) was added to the histones. The samples were vortexed for 2 minutes at room temperature. To quench the derivatization reaction, 1 µL of 80 mM hydroxylamine was added and the samples were incubated for 20 minutes at room temperature. The pH of the samples was measured to ensure optimal trypsin digestion, ~pH 8.5. 250 ng of trypsin was added to each sample and left to digest at 37°C for 4 hours or overnight. Upon digestion, the second round of derivatization was carried out by adding 3 µL of 1:100 diluted propionic anhydride dissolved in acetonitrile. The samples were then vortexed for 2 minutes at room temperature. 3 µL of 80 mM hydroxylamine was then added to quench the second round of derivatisation.

2.8.2. In-gel derivatisation

Modified from Bonaldi et al (2004) (Bonaldi et al. 2004). SDS-PAGE gel was prepared as previously stated. However, the protein was only allowed to migrate from the stacking into the running gel. Once all of the desired proteins had migrated, by checking the corresponding band on the protein ladder, the electrophoresis was stopped and the gel

was stained with coomassie (InstantBlue Protein Stain) for at least an hour. The gel was then washed with ddH₂O. The stained section was cut out from the gel with a scalpel. The gel piece was cut into three pieces: not so small as to get aspirated by a pipette but small enough to allow for a larger exposed surface area. The gel slices were then shaken for three 5 minutes washes in destaining/dehydrating solution: 50% acetonitrile, 100 mM TEAB. Any residual liquid was then removed. The gel pieces were dehydrated, and so were rehydrated with 100 mM TEAB. The gel pieces were submerged with a 75:25 acetonitrile to propionic anhydride mixture that was prepared immediately before being added to the gel pieces. The gel pieces were left to shake at room temperature for 30 minutes. The liquid was then discarded, and the gel pieces were washed twice in 100 mM TEAB with 5-minute vigorous shaking.

The gel pieces were then dehydrated as before, two changes of destaining/dehydrating solution, with 5 minutes shaking between each change. The liquid was removed. The gel pieces were then spun in a vacuum centrifuge (Savant™ SPD131DDA SpeedVac™ Concentrator) without heating. Dilute trypsin solution was prepared: 25 ng/μl, in 100 mM TEAB. The gel pieces were rehydrated in trypsin solution so that the gel pieces are full size. The gel pieces were left on ice for 5 minutes to swell up. Excess trypsin solution was removed. The gel pieces were submerged in 100 mM TEAB. They were incubated for at 37 °C for four hours or overnight. An hour after incubation, the gels were monitored to assess if more 100 mM TEAB solution was required.

The next day, the gel pieces were vortexed, spun down and the supernatant that contains digested peptides were collected. Acetonitrile equivalent to the volume of gel pieces was

added to the tubes. They were then shaken for 5 minutes on the vortex. The supernatant was then removed and added to the other peptide containing supernatant. This acetonitrile dehydration was repeated twice more to dehydrate the gel pieces. The supernatant was then spun in the vacuum centrifuge to concentrate the solution to 10 μ L. The peptides were derivatized again following the in-solution derivatization protocol.

2.8.3. In solution trypsin digestion of peptides

5 μ g of prepared histone solution (either with acid extraction or Mnase digestion) were made up to 10 μ L with 100 mM TEAB. 250 ng of trypsin (Sequencing Grade Modified Trypsin Promega Cat: V5111). Samples were incubated at 37 °C for four hours or overnight. Upon digestion, the samples were acidified by adding trifluoroacetic acid (TFA) to give a final concentration 0.1% TFA.

2.8.4. On-bead Trypsinisation

Upon completion of IP washes, the magnetic beads were washed twice with 1 mL of cold 100 mM TEAB solution, which was freshly made and kept at 4 °C. After the second TEAB wash, the magnetic beads were transferred to new tubes.

10 μ l (10 ng/ μ l) of trypsin in 100 mM TEAB, was added directly to the washed beads, to digest the bead-bound protein. An enzyme-to-protein ratio of 1:100 (wt/wt) was used. The samples were vortexed for 15 seconds every 3 minutes for the first 15 minutes to ensure that the beads were evenly suspended in the trypsin solution. The mixture was then left to digest overnight at 37 °C in an oven incubator without further agitation of the beads.

After the overnight digest, 10 µL of trypsin, at the same previous enzyme-to-protein ratio, to each sample was added and left to digest for 4 hours at 37 °C. The tubes were placed back on the magnetic rack. The supernatant, which now contains the digested peptides was removed and retained. The samples were acidified by adding TFA to the solution to give a final concentration of 0.5% TFA.

2.9. Peptide Desalting and Detergent Removal

2.9.1. Stage-Spin column desalting

Modified version from Rappsilber et al. (2007). Desalting tips were prepared by punching out a circle of Empore C18 membrane with a blunt-ended 250 µL syringe. Fused-silica was used in place of the plunger to push out the C18 membrane into a 20 µL pipette tip. Three circles were pushed into the pipette tip as far as they can fit. Used miniprep columns from QIAGEN QIAprep Spin Miniprep Kit were used as a bracket for the tips containing the C18 membranes. The silica from miniprep was hollowed out and discarded. When the column was placed into a 1.5 mL microcentrifuge and the C18 membrane tip was placed into the column, the bottom of the tip would not come into contact with the bottom of the microcentrifuge tube. In 1.5 mL microcentrifuge tube, two solutions were prepared:

- Solution A: 80% acetonitrile, 0.1% TFA.
- Solution B: 0.1% TFA.

The C18 membrane was wetted by adding 20 µL Solution A to the top of the tip. They were spun for 1 minute at 2000 rcf on table-top microcentrifuge. Once confirmation that all of Solution A had passed through the C18 membrane, the solution in the microcentrifuge tube was discarded. If the whole solution had not passed through the membrane, then another spin session on the microcentrifuge was applied. The check was

applied after every microcentrifuge spin. 20 μ L Solution B was added to the C18 membrane to equilibrate the membrane and then spun at 2000 rcf for 1 minute.

The sample was acidified (0.1% TFA) and then loaded onto the tip. The samples were spun through the membrane for five minutes at 2,000 rcf. 40 μ L Solution B was loaded on to the C18 membrane tip to wash and spun at 3.5 minutes at 200 rcf. The tips were then transferred to newly labelled microcentrifuge tubes. 20 μ L solution A was loaded, spun for 3 minutes at 2000 rcf. A final check to ensure all of the liquid has passed through the C18 membrane. The samples were spun on the vacuum concentrator with no heat to evaporate the acetonitrile. After acetonitrile evaporation, 4 μ L of the sample remained. The volume was increased to 20 μ L by adding 0.1% TFA.

2.9.2. Sera-Mag CM clean up of peptides

A stock solution of the Sera-Mag beads was prepared (DSMG-CM 1 μ M 5% solids: Fisher part #11819912). Beads were taken out of the fridge for 10 minutes. 20 μ L of Sera-Mag beads A were added to 20 μ L of Sera-Mag beads B. 160 μ L of water was added to the mixture. The mixture was placed on a magnetic rack and the supernatant was discarded. The tubes were taken off the magnetic stand and the beads were then rinsed with 200 μ L of ddH₂O and mixed by pipetting up and down repeatedly. Two more identical washes were carried out. The beads were stored in 100 μ L of water in the fridge.

The stock beads were vortexed for a minute, and 2 μ L (20 μ g) of beads were added to each sample of prepared histones (20-100 μ L volume). 5 μ L of neat formic acid was added to 995 μ L acetonitrile (FA/MeCN). An equivalent volume of FA/MeCN were added to the

prepared histone sample to give a 50% FA/MeCN concentration. The samples were incubated upright at room temperature for 8 minutes. The tubes were then placed on a magnetic rack for 2 minutes. The supernatant was carefully removed and discarded. The beads were rinsed with 200 μ L 70% ethanol. They were left to incubate for 30 seconds and then placed on the magnetic rack for 2 minutes. The supernatant was then removed and discarded. Care was taken to ensure no Sera-Mag beads were aspirated. The protein was eluted in 20 μ L of 100 mM TEAB, the beads were mixed by pipetting up and down, and then left to incubate for 5 minutes.

2 μ L of 1% (v/v) propionic acid (Fisher Scientific: A258-500) in acetonitrile was added to the solution and then vortexed for 2 minutes at room temperature. 2 μ L of 80 mM hydroxylamine was then added to quench the propionylation reaction. The pH was measured to ensure it was in the range of pH 7.5-8.5. Trypsin was added to the solution at a 1:20 protein ratio (wt/wt). The samples were left to digest overnight. Upon digestion, acetonitrile was added to reach 87-97% final concentration. The sample was left to incubate for 8 minutes and then the supernatant was removed and discarded. The peptides were eluted for the second round of derivatization by adding 20 μ L of 100 mM TEAB and left to incubate with slight agitation by vortexing and pipette mixing for 5 minutes at room temperature. 2 μ L of freshly prepared 1% propionic anhydride in acetonitrile was added to the solution and then vortexed for 2 minutes at room temperature. 2 μ L of 80 mM hydroxylamine was then added to quench the reaction. The samples were incubated upright for 20 minutes at room temperature.

Formic acid was added to the samples for a final concentration of 5% so that the pH was acidic. Acetonitrile was then added to reach 87.5-95% final concentration. The samples were incubated for 8 minutes at room temperature and then the supernatant was discarded using the magnetic stand. The peptides were eluted in 2% DMSO and then sonicated for 1 minute using the diagenode Biorupter Pico. The tubes were then spun down at maximum speed for 1 minute and then placed on the magnetic stand. The peptide-containing supernatant was transferred to a fresh microcentrifuge tube. TFA was added to a final concentration of 0.1%.

2.10. Phosphoenrichment

Peptide samples were resuspended in Loading Buffer (80% acetonitrile, 6% trifluoroacetic Acid). A constricted GELoader tip was prepared as previously described using Empore C8 filter membrane material by punching out a small C8 disk using the tapered end of a 1000 μ L pipette tip. Circular movements were applied to rotate the tip to ensure complete excision of the disk. The 1000 μ L tip was placed into a 200 μ L tip and fused-silica tubing was used as a plunger to fix the disk into the tapering of the GELoader tip. The GELoader tip was cut at the tapered end, approximately 0.2 cm away from the C8 plug. TiO₂ beads (Titansphere; 5 μ m; GL Sciences, Japan) were resuspended at 10 mg/mL in 30% ACN and 0.1% TFA. The C8 constricted tip was washed with 20 μ L of methanol by centrifugation (100 g, 1 minute, RT). 50 μ L of the TiO₂ slurry was packed into the prepared GELoader Tip-Column by centrifugation (100 g for 5 minutes at RT). The TiO₂-C8 microcolumn was conditioned with 50 μ L Loading Buffer (150 g for 5 minutes at RT). The samples were then loaded onto the microcolumns by centrifugation (50 g for 30 minutes at RT). The microcolumns were washed with: 50 μ L Wash 1 Buffer (50% ACN, 0.5% TFA, 0.2 M NaCl) at 150 g for 20 minutes at RT. A second wash was carried out with 50 μ L Wash 2 Buffer (50% ACN, 0.5% TFA): at 150 g for 20 minutes at RT. The outside tip was wiped using ethanol-wetted Kimtech wipes. The microcolumn was transferred to a new pipette tube. 35 μ L of 10%

(v/v) formic acid (FA) was pipetted in to a new tube. The sample was eluted into the same tube with two elutions: 20 μ L Elution Buffer 1 (10% NH_4OH): 100 g for 20 minutes at RT and then with 3 μ L Elution Buffer 2 (80% ACN, 2% FA): 100 g for 20 minutes at RT. 3 μ L of 100% FA was added to further acidify the samples. The samples were vortexed and briefly spun down. Samples were dried down using a Thermo SpeedVac and resuspended in 0.1% FA.

2.11. Nano-LC/MS

ThermoFisher U3000 nanoLC and Orbitrap XL mass spectrometer were used to analyse peptide samples. The nanoLC was equipped with a Pepmap100 C18 trapping cartridge. Peptide samples were loaded in 0.1% TFA loading buffer at a flow-rate of 5 μ L/min for five minutes. Peptides separation was conducted with a PepMap100 analytical column in 0.1% FA at a flow rate of 0.3 μ L/min; a 1-35% gradient with increasing acetonitrile is applied, over 28 minutes.

Time (min)	0	15	25	28	28	35	35
Acetonitrile %	1	12	35	60	90	90	1

Electrospray Ionisation was performed using a New Objective nanospray emitter with a 10 μ m tip (FS360-20-10-N-20). The linear ion trap of Orbitrap XL was used to conduct Pseudo-SRM with the following settings:

- Precursor Isolation Window: 2 m/z
- Ion-trap fill time: maximum of 50 ms
- Target Ion Count: 1×10^4

Orbitrap was used to carry out high resolution precursor scan with a total cycle of <30 s which provided enough points across peptide peaks for quantification. A data-dependent

event was performed concurrently with targeted LTQ events on TiO₂-enriched eluates to identify untargeted enrichment.

2.12. Data Analysis

Skyline v3.5 (MacCoss Lab, University of Washington) (MacLean et al. 2010) was used for the development of pseudo-SRM methods and for analyzing data generated from LC-MS/MS. Sequences for peptides corresponding to targeted proteins were downloaded from Uniprot (Uniprot Consortium, 2017) and imported into Skyline. The peptides were modified in-silico with reported and speculative post-translational modifications. Possible permutations of m/z were generated after in-silico digestion with trypsin and chemical derivatisation. To ensure we were detecting different peptides that were isobaric, we monitored for b and y ions that would differentiate between the isobaric peptides.

LC-MS/MS data was imported into Skyline and exported as CSV file into Microsoft Excel. Precursor ion signal was deducted from total ion signal to obtain fragment ion signal for the peptides. Fragment ion signal corresponding to the same peptide together (i.e. different modifications on the same sequence) were combined to get total fragment signal. Peptide of interests were quantified by normalising their signal intensity against a series of internal control peptides (listed below) that did not carry any known modifications and were available in all known isoforms of Histone H3 and Histone H4. Therefore, the control peptides would have similar levels of abundance across the different samples. Error bars were generated using standard deviation across different biological replicates. Triplicate data was generated from three different biological samples. For γ H2AX-IP-MS, two-tailed T-tests were carried out on all time points eluates against undamaged global values and corresponding global time point.

Table 2.1 H2AX Control Peptides Used for Quantifications

Histone	Sites	Sequence	Monoisotopic Mass (Da)	Derivatised
Histone H2AX	39 to 43	GHYAER	731.324	No
Histone H2AX	135 to 142	ATQASQEY	896.377	No
Histone H2AX	135 to 142	ATQASQEY	976.343	No
Histone H2AX	135 to 142	ATQASQEY	976.343	No

Table 2.2 Control Peptides Used for Quantifications of pSRM

Histone	Sites	Sequence	Monoisotopic Mass (Da)	Derivatised
Histone H3	41 to 49	Y[+56]RPGTVALR	1,089.628	Yes
Histone H3	57 to 63	S[+56]TELLIR	888.527	Yes
Histone H4	68 to 78	D[+56]AVTYTEHAK[+56]R	1,403.703	Yes
Histone H4	46 to 55	ISGLIYEETR	1,181.628	No
Histone H3	41 to 49	YRPGTVALR	1,033.602	No
Histone H4	24 to 35	DNIQGITKPAIR	1,326.761	No

Table 2.3 Synthetic peptides and recombinant histones

Peptides listed below were synthesised by JPT (Germany)

Synthetic Peptide Sequence
GHYAERATQASQEY
GHYAERATQA-S[p]-QEY
GHYAERATQASQE-Y[p]
KGHYAERGKTGGKAR
KGNYAERGKTGGKAR
GKTGGKARAKAKSR
GKQGGKARAKAKSR

Synthetic peptide sequence locations are in bold below:

H2AX:

MSGR**GKTGGKARAKAKSR**SSRAGLQFPVGRVHLLR**KGHYAER**VGAGAPVYLAADVLEYLTAE
ILELAGNAARDNKKTRIIPRHLQLAIRNDEELNKLGGVTIAQGGVLPNIQAVLLPKKTSATVGPKAP
SGGKK**ATQASQEY**

H2A1B:

MSGR**GKQGGKARAKAKTR**SSRAGLQFPVGRVHLLRKGNYSERVGAGAPVYLAADVLEYLTAEIL
ELAGNAARDNK KTRIIPRHLQLAIRNDEELNKLGRVTIAQGGVLPNIQAVLLPKKTESHHKAKGK

Heavy labelled recombinant H2AX and H2A1B were synthesised in collaboration with Andrew Flaus group, National University of Ireland, Galway (Hatimy, et al 2015).

Chapter 3: Global H2AX Analysis

3.1. Introduction

3.1.1. H2AX

In regards to variants, the H2A family of histones is the most diverse and numerous. Structurally, the C-terminus of H2A variants are located at the point where the DNA double helix enters and exits the nucleosomes (Bonisch & Hake 2012). By possessing such a large diversity of amino-acid sequence at the C-terminus, there is an increased assortment of intra-chromatin interactions. Compared to the other histone H2A variants, H2AX is unique in that it possesses a distinctive 22 amino acid C-terminal (Pinto and Flaus 2010). While the sequence can vary slightly between closely related species, H2AX possesses a conserved characteristic motif, SQ[E/D]* (* being a hydrophobic residue) that differentiates it from other H2A variants (Bonisch & Hake 2012). Upon DNA damage, Serine 139 of H2AX is phosphorylated by one of the PIKK kinases. Essentially, H2AX is crucial for maintaining genome stability and integrity, and not just in the context of DNA damage.

There are few “universal” H2A variants, that can be found in almost all organisms. H2AX and H2AZ are two examples of that can be found in almost all higher eukaryotes including plants (Talbert & Henikoff 2010). Even if the variant is not present in a given species, the phosphorylation of H2A at the C-terminal end at the onset of DNA damage is a conserved phenomenon. Both *S. pombe* (Nakamura et al. 2004) and *S. cerevisiae* (Downs et al. 2000) possess the SQ[E/D]* motif in their H2A variants, the consensus site for PIKK kinases. When DNA damage takes place, S129 of yeast is phosphorylated similarly to S139 of H2AX. Mec1 (ATR) and Tel1 (ATM) are the kinases responsible for phosphorylating H2A

S129 in the presence of DNA damage. Once S129 is phosphorylated, Rhp9(Chk1) is recruited showing that it is a conserved process.

It has also been observed that when the C-terminus of H2AX is phosphorylated, the weak affinity between histone H1 and H2AX is weakened further (A. Li et al. 2010); the phosphorylation modification makes H2AX even more negatively charged resulting in DNA having reduced binding with the nucleosome, making the chromatin more accessible for DDR machinery.

Even though H2AX and its phosphorylated form, γ H2AX, is seen to facilitate the DDR it is not essential for DSB repair. Cells are still viable. However, H2AX^{-/-} mice are much more sensitive to ionising radiation compared to their H2AX^{+/+} counterparts. H2AX^{-/-} mice also possess increased chromatid breaks and dicentric chromosomes. (Celeste, Petersen et al. 2002).

Another function of H2AX not thought to be related to the DNA damage response directly is during oocyte development. Upon fertilisation in mammals, the protamines from the sperm are rapidly replaced by histones. Histones H3.3 and H2AX were deposited into the paternal chromatin. The particular conformation is thought to be required as a marker for maternal cell machinery to identify which parental chromatin to strip DNA methylation from. Interestingly, it was not the phosphorylated S139 that was important in this case but the C-terminal end of H2AX (Nashun et al. 2015).

The modifications of H2AX that we shall focus on in this section are S139phos and Y142phos. Even though phosphorylation of Threonine 136 has been reported to occur in

mammalian systems (A. Li et al. 2010), the precise function of T136phos has yet to be elucidated. It is theorised that upon DNA damage, S139phos along with T136phos can modify chromatin structure (A. Li et al. 2010).

3.1.2. S139 phosphorylation

In quick response to DNA damage, the C-terminal domain of H2AX at S139 is phosphorylated (Rogakou et al. 1998). When a double strand break has occurred, MRE11, RAD50 and NBS1 bind to the double-strand break and in turn recruit one of ATM, ATR and DNA-PK, which are the kinases responsible for phosphorylating S139 (Podhorecka et al. 2010). All three kinases are phosphatidylinositol 3-kinase-related kinases. Due to its rapid response to DNA damage and ease of detectability, the presence of γ H2AX is used as a diagnostic marker of DSB (Heylmann & Kaina 2016). Phosphorylation of H2AX S139 can take place within 1 minute of DNA damage taking place. The spread of the phosphorylation can spread megabases away from the site of damage in a bi-directional manner. Rokagou *et al.* saw that for every Gy of ionising radiation, you get an increase of 1% of total H2AX becoming γ H2AX. Rokagou *et al.* also deduced that one double strand break would result in γ H2AX spreading so that its coverage across the genome encompasses 2 million base pairs (Rogakou, Pilch et al. 1998). To say there is only one exact biological function of megabase-wide amplification of γ H2AX would be contentious with the overwhelming body of evidence that has specified different functions of γ H2AX in the context of DDR alone.

Perturbed recruitment of 53BP1, NBS11, and BRCA1 to IRIFs was demonstrated in cells that had H2AX knocked out (Celeste et al. 2002). There were increased chromosomal aberrations in mouse thymocytes that had the genotypes H2AX^{+/-} and H2AX^{-/-} (Celeste, Difilippantonio, et al. 2003), (Petersen et al. 2001) and (Bassing et al. 2002). It was also observed that mouse embryonic stem cells lacking phosphorylated H2AX had an increased sensitivity to DNA damage as both the NHEJ and HR pathways were demonstrated to be sluggish in its efficacy in repairing breaks (Bassing et al. 2002). Also, H2AX knockout mice were both sterile and possessed compromised in class switch recombination as there was a reduction in secondary immunoglobulin isotypes levels. At the immunoglobulin locus, chromosomal abnormalities were observed; as the cell induces DSBs during CSR, the cell was then not able to repair the breaks in an efficient manner (Petersen et al. 2001).

The culmination of these studies indicate that γH2AX is a docking site for DDR proteins and facilitates the DDR: phosphorylation of H2AX at the C-terminal tail, especially S139phos, acts an epigenetic signal that acts as a signal for a handful of proteins that possess inherent binding sites specific for the phosphorylated H2AX C-terminal tail.

3.1.3. H2AX Kinase

ATM is in an inactive state when there is no DNA damage. It forms a homodimer with another ATM molecule. At the onset of DNA damage, ATM phosphorylates itself at serine 1981, and the ATM pair disassociate from each other (Rogakou, Pilch et al. 1998) (Helt et al. 2005). Upon chromatin remodelling, NBS1 recruits ATM to the site of damage. ATM proceeds to phosphorylate both checkpoint proteins and downstream DDR proteins (Lou et al. 2006). It was demonstrated that ATM recruitment to DSB is hindered in the absence of γ H2AX. Any DDR pathway downstream that takes place as a result of ATM activation via ATM recruitment to the site of damage would be affected as a result. (Kobayashi et al. 2004).

3.1.4. MDC1

Knocking out MDC1 (Mediator of DNA damage checkpoint protein 1) in mice exhibited a phenotype similar to that of H2AX being knocked out: Male mice are infertile, and mice exhibit slowed growth. H2AX S139 phosphorylation is regulated by MDC1. MDC1 also mediates the interaction between ATM and γ H2AX. MDC1 interacts with H2AX through its BRCT domain. Also, ATM interacts with MDC1 with the FHA domain on MDC1; speeding up ATM-dependent recruitment to the site of damage (Lou et al. 2006). Stewart et al. used human cell extract to demonstrate that MDC1 interacted with phosphopeptides during phosphor-enrichment (Stewart et al. 2003). MRN was also shown to interact with MDC1. Shortly after, it was determined that MDC1 BRCT domain recognises the H2AX C-terminal domain uniquely (Goldberg et al. 2003). In vivo experiments showed that there is a significant reduction in recruitment of MDC1 to IRIFs when the conserved residues in the BRCT domain are mutated. MDC1 also had reduced recruitment to IRIFs when either Serine 139 or Tyrosine 142 of H2AX were mutated. As a result, cells were much more sensitive to irradiation (Stucki et al. 2005).

3.1.5. 53BP1

In the presence of a DSB, chromatin-bound ATM will phosphorylate MDC1. In turn, MDC1 will recruit RNF8 and RNF168. RNF168 will ubiquitinate histone H2AK13/K15, including H2AX (B. Wang & Elledge 2007; Mattioli et al. 2012). 53BP1 will be recruited to the site of damage by interacting with H2AK13ub/K15ub with its ubiquitination-dependent recruitment (UDR) motif. The Tudor domain of 53BP1 will then interact with H4K20me1/me2. 53BP1 cannot dock in the absence of either one of H2AK13/15ub and H4K20me1/me2. Even though 53BP1 possesses a BRCT domain, it is dispensable for recruitment to damage sites. 53BP1 activates ATM by interacting with MRN through RAD50 and its BRCT domain. In turn, more ATM is recruited to the site of damage (Lee & Paull 2007).

Ward et al., demonstrated that when cells are deficient in H2AX, there is reduced 53BP1 foci formation implying H2AX plays a role in 53BP1 recruitment (Ward et al. 2003). More recently, 53BP1 has been shown to bind to γ H2AX S139 in both in-vitro and in-vivo. (Kleiner et al. 2015) (Baldock et al. 2015) 53BP1 binding to γ H2AX with its BRCT is required for ATM foci formation at the site of damage. Baldock et al. also showed that a version of the 53BP1 void of BRCT₂ domains was not able to recruit NBS1 or phosphorylated ATM. If 53BP1 has a mutated BRCT domain perturbing the binding to γ H2AX, NBS1 could still be recruited (and without the MRN complex) to the site of damage but could not recruit pATM (Baldock et al. 2015). Lastly, they demonstrate if pATM could not be recruited as a consequence of the cells lacking functional 53BP1, the number of γ H2AX foci in the late time point would be significantly higher than that of cells that were mock treated. Consistent with previous work showing that in the absence of ATM, breaks in heterochromatic regions persist at later time points (Goodarzi et al. 2008).

3.1.6. Y142 phosphorylation

The Williams Syndrome Transcription Factor (WSTF) remodelling factor kinase is the only known kinase in somatic cells responsible for phosphorylating tyrosine 142 on the H2AX C-terminal (Cook, Ju et al. 2009; Krishnan, Jeong et al. 2009; Xiao, Li et al. 2009). In the presence of DNA damage, H2AX Y142 within the vicinity of the damage is dephosphorylated so as to avoid an unnecessary apoptotic cell fate (Cook, Ju et al. 2009). Cook et al. demonstrate binding of proteins involved in apoptosis with doubly-phosphorylated H2AX S139 Y142 synthetic peptides. Compared to S139 phosphorylation, Y142 phosphorylation is not as well documented. However, the kinases and phosphatases responsible for modifying the residue have been characterised, but accurate basal levels of Y142 phosphorylation have yet to be attained. Previous estimates suggest that the modification cannot possess a basal level higher than 10% (Tran, Zamdborg et al. 2011). Xie et al. transfected HA-tagged H2AX into H2AX deficient mouse cell line. They irradiated their cells with 50 Gy of IR, and they were able to detect six post-translational modifications of H2AX, but Y142 phosphorylation was not amongst the modifications detected using the LC-MS screen. However, as they did not specify the recovery time allotted to the cells after irradiating them with 50 Gy of ionising radiation, it is difficult to say whether Y142 phosphorylation corresponding to apoptotic signals as a response to damage would have been activated (Xie, Odate et al. 2010). Another role

The part that Y142 phosphorylated plays in the DDR is of great importance. It would also be interesting if interacting partners could be identified of the doubly phosphorylated

isoform. (Singh, Basnet et al. 2012). Brown et al., identified that by mutating Y142A mutation alone exhibited sensitivity to DNA damage. However, when both S139A and Y142A mutation were present then the IR sensitivity was rescued (Brown, Eykelenboom et al. 2012); implicating Y142 in both apoptosis and DDR depending on if the residue is phosphorylated or not, respectively.

3.1.7. H2AX levels across cell lines and in the genome

γ H2AX plays a vital role in the DDR; it is then paramount to see whether different cell lines express the gene for H2AX in varying amounts or at a more homogenous level. If the levels vary between different cell lines, would the discrepancy have leading effects on the response time or intensity to damage? Previous studies have shown that H2AX levels can vary in different cell lines from as low as 2% to as high as 25% of the total H2A (Rogakou, Pilch et al. 1998).

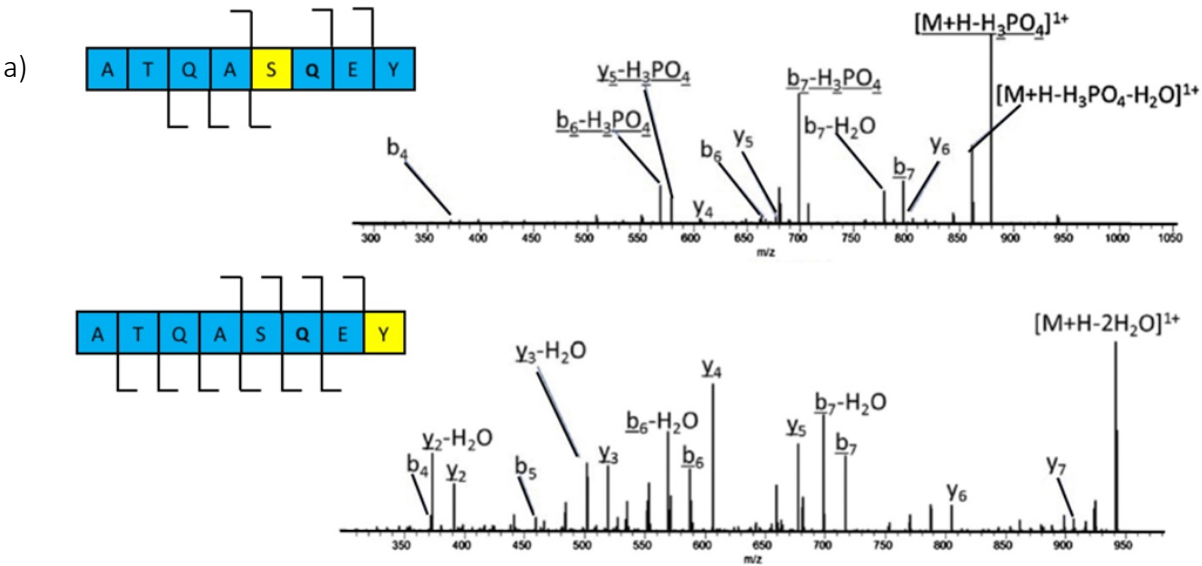
3.2. Results

3.2.1. S139 phosphorylation

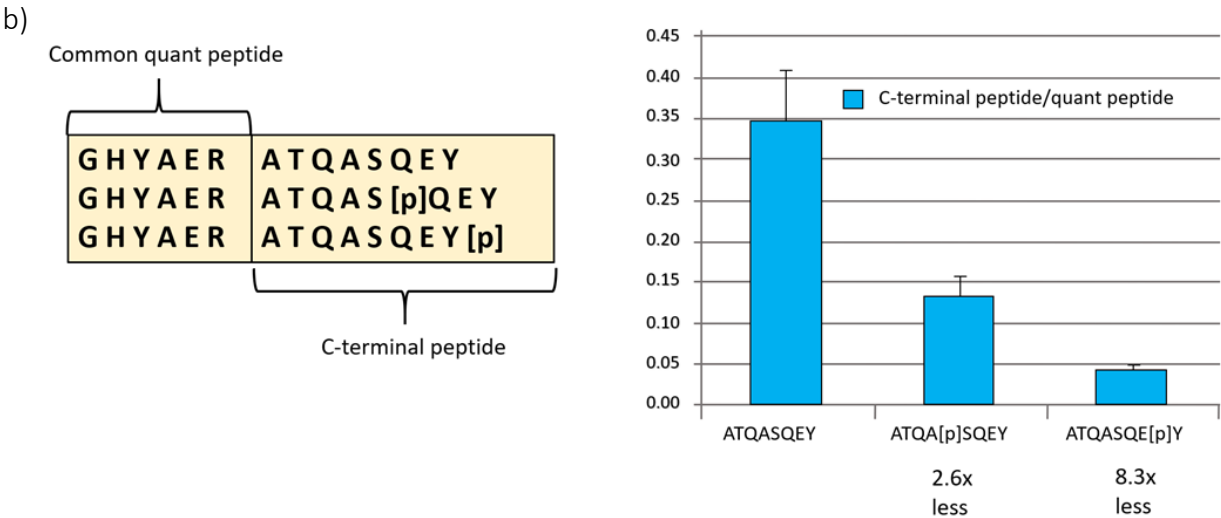
Synthetic peptides corresponding to the C-terminal of H2AX were designed for H2AX phosphopeptides SRM assay (Methods and Materials). Three different synthetic peptides were developed; the phosphorylated S139 residue, phosphorylated Y142 residue or no phosphorylated residues. To determine how the three different synthetic peptides would ionise compared to each other, all three peptides were concatenated with an identical quantification peptide. A simple trypsin digest would allow for both the control peptides to separate from the peptides of interest. Using the linear ion trap of an Orbitrap XL, the C-terminal synthetic peptides were fragmented (**Figure 3.1a**) and their ionisation efficiencies were quantified with pseudo-SRM assays. Singly and doubly-charged ions were detected for the H2AX C-terminal peptides. The m/z region corresponding to the singly charged peptide exhibited low background. In the presence of abundant canonical H2A peptides, the C-terminal peptide may be harder to detect in double-charged form. The common quantification peptide released upon trypsin digestion acts as an internal control because the ionisation efficiency of GHYAER (**Figure 3.1b LHS**) should not be affected by the phosphorylated or lack of connection to the other peptide. Since the ionisation efficiency should be equal for the internal common quantification peptide, it can be used as a standard for the C-terminal peptide to be measured against. In **Figure 3.1b RHS** we see that the phosphorylated peptides were detected less compared to the unphosphorylated form. Phosphorylated S139 and phosphorylated Y142 peptides were detected 2.6x less and 8.3x less than the unphosphorylated peptide, respectively. The new values could be applied to correct data corresponding to percentages of H2AX that

possess phosphorylated S139 or Y142 in cells that have been treated with ionising radiation.

Figure 3.1. MS/MS Fragmentation of ATQASQEY S139 orY142 phosphorylated precursors



Synthetic Peptides



c)

Synthetic Peptide Sequence		
GHYAERATQASQEY	KGHYAERGKTGGKAR	GKTGGKARAKAKSR
GHYAERATQA-Sp-QEY	KGNYAERGKTGGKAR	GKQGGKARAKAKSR
GHYAERATQASQE-Yp-		

3.1 MS/MS Fragmentation and ionization efficiency of ATQASQEY S139 and Y142 singly phosphorylated precursors

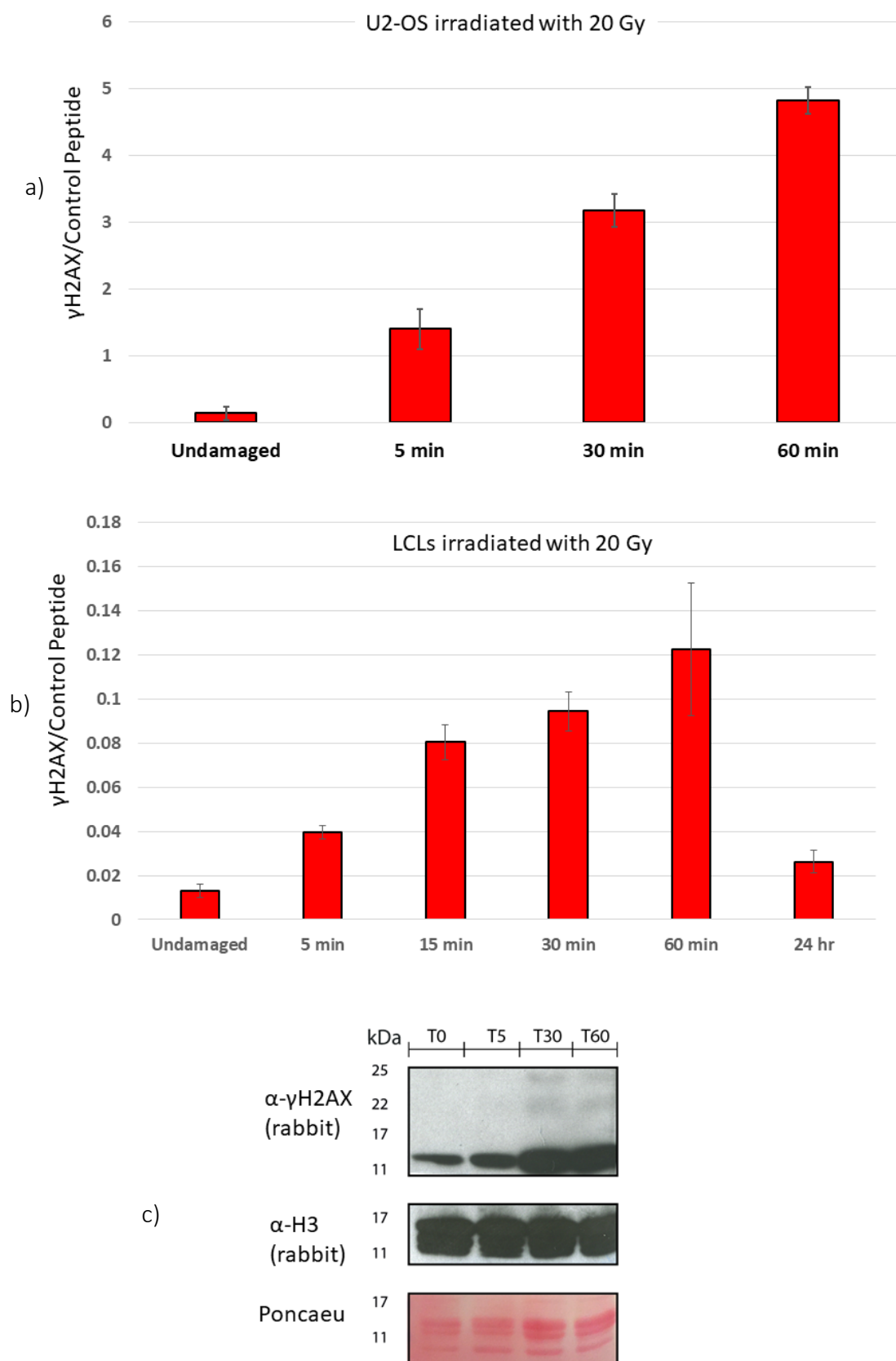
a) MS/MS fragmentation of singly charged precursors for residues corresponding to phosphorylated S139 or Y142 from the synthetic peptide sequence ATQASQEY. Fragment ions detected are highlighted on the left hand side. Fragment ions used for quantitation are underlined.

b) Ionization efficiency of differently modified C-terminal peptides. LHS: Synthetic peptides were designed concatenated to a common quantitation peptide, to be released by trypsin digestion. RHS: quantification of ionization efficiency after trypsin digestion, nanoLC separation and pseudo-SRM peptide fragmentation. Error bars correspond to standard deviation, n = 3 separate digestions.

c) Table of synthetic peptide sequence used in this assay. Peptides listed in the table were synthesized by JPT (Germany).

Applying the correction factor results in an inflated values of γ H2AX levels compared to regular H2AX. In **Figure 3.2** we see that γ H2AX levels peaked at around 30 minutes for both U2-OS cells and LCLs. This implies that the response of DDR is consistent between the two cell lines in that γ H2AX reaches its maximum level at the same time in both cell lines. If applied to γ H2AX levels obtained from irradiated LCL and the U2-OS cells we observe levels reaching 5-10 times higher than undamaged cells after being subjected to 20 Gy of ionising radiation.

Figure 3.2. Quantification of global γ H2AX S139 phosphorylation after DNA damage



3.2. Quantification of global γ H2AX S139 phosphorylation after DNA damage

- a) Time course after 20 Gy IR treatment in U2-OS. Fragment ions for singly charged ATQASQEY and ATQAS[p]QEY precursors were quantified against internal control peptides. Error bars correspond to standard deviation; n = 2 biological replicates. Technical triplicate were used to analyse each of the two biological replicates.
- b) Time course after 20 Gy IR treatment in LCLs. Fragment ions for singly charged ATQASQEY and ATQAS[p]QEY precursors were quantified against internal control peptides. Error bars correspond to standard deviation. Technical triplicate were used to analyse each of the three biological replicates.
- c) Western blot of α - γ H2AX after 20 Gy IR treatment in U2-OS. Immunoblot Histone H3 and Ponceau staining of the membrane post-transfer were used to control for loading.

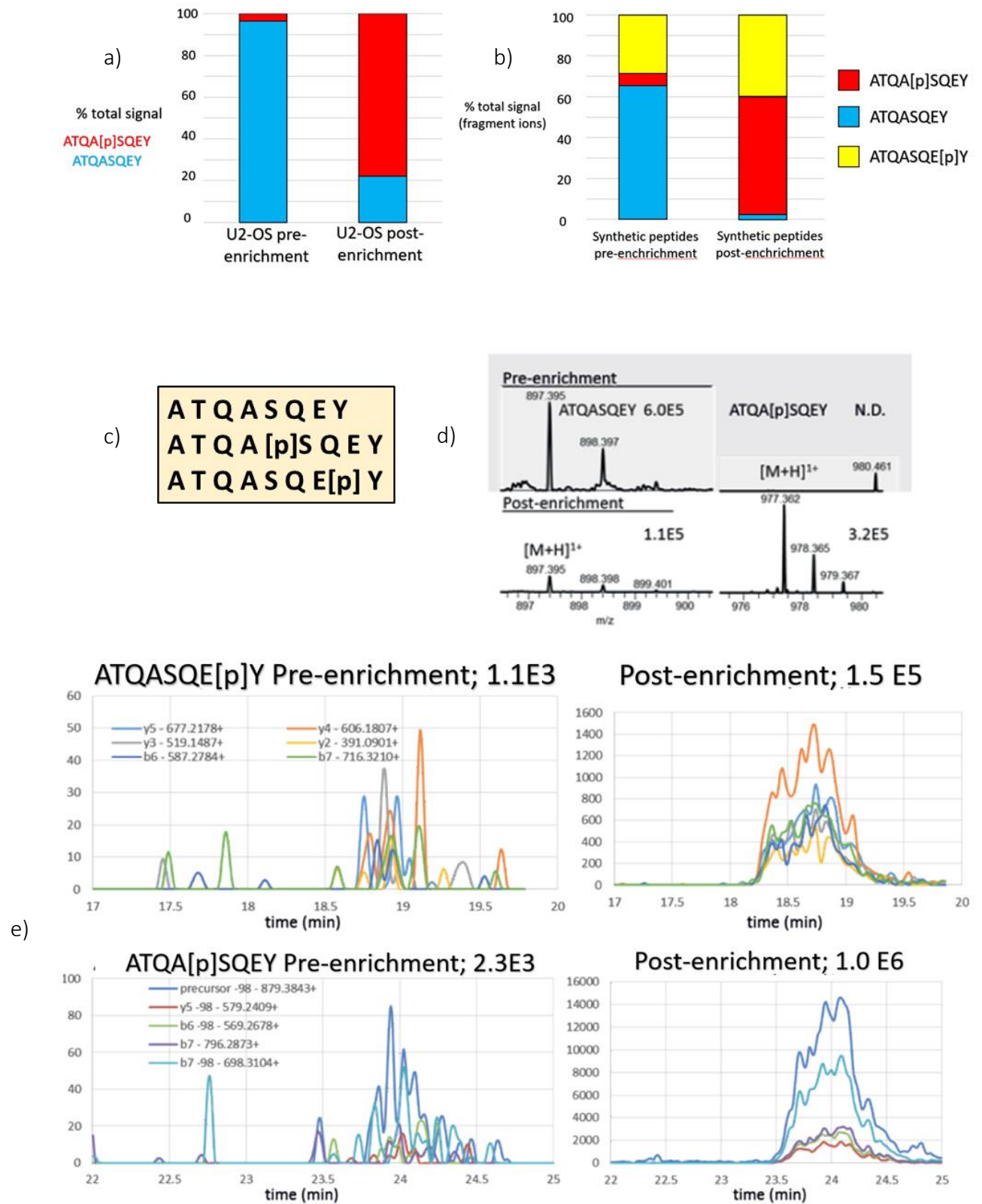
3.2.2. Y142 phosphorylation

While quantifying S139 phosphorylation in U2-OS cells, Y142 phosphorylation was also targeted. Y142 phosphorylation could not be detected in the undamaged U2-OS cells nor in the post-IR recovery time points shown in **Figure 3.2a**. As the literature states that Y142 phosphorylation is most prominent in the absence of DNA damage, a TiO₂ metal oxide affinity phosphoenrichment was carried out on histones prepared from acid extracts corresponding to U2-OS cells that were not subjected to IR. Post-enrichment, these cells were shown to have a low percentage of S139 phosphorylation, approximately 4%, as seen in **Figure 3.2**. Alongside, the phosphoenrichment from U2-OS acid extracted prepared histones a control sample was also used that was a cocktail of 50 pmol BSA, 0.5 pmol α -casein, phosphorylated S139 synthetic peptide, and phosphorylated synthetic Y142 peptide. In the U2-OS post-enriched samples, levels of γ H2AX were elevated similar to that of damaged cells (**Figure 3.3a**). Also, other phosphorylated peptides were detected in the post-enriched samples, but no phosphorylated Y142 peptide was detected from the U2-OS post-enriched sample. In the control sample, we detect increased enrichment of both phosphorylated S139 phosphorylated Y142 peptides, which demonstrated that there was no technical fault with the enrichment process.

Another control to ensure that there was no fault with the technique was applied: 2 pmol of the phosphorylated Y142 synthetic peptide was spiked into a 50 μ g sample of undamaged of U2-OS acid extract. Enrichment was carried out in parallel with an unspiked equivalent sample. As shown in **Figure 3.3a**, γ H2AX was enriched in both the spiked and unspiked sample (**Figure 3.3d**). In **Figure 3.3d**, we see that phosphorylated Y142 was enriched in the spiked sample and compared to pre-enrichment the signal was much more pronounced. Even in the presence of a histone rich environment, the TiO₂

enrichment allowed for the enrichment of 232 other distinct phosphopeptides from non-histone proteins such as CDK3 Y15 phosphorylation, see **Appendix 1**. Histone phosphopeptides included H3 S10, H3 S57, H4 S47 and nine phosphorylated peptides from histone H1.

Figure 3.3. Y142 Phosphorylation



3.3 H2AX Y142 Phosphorylation is a low abundance modification

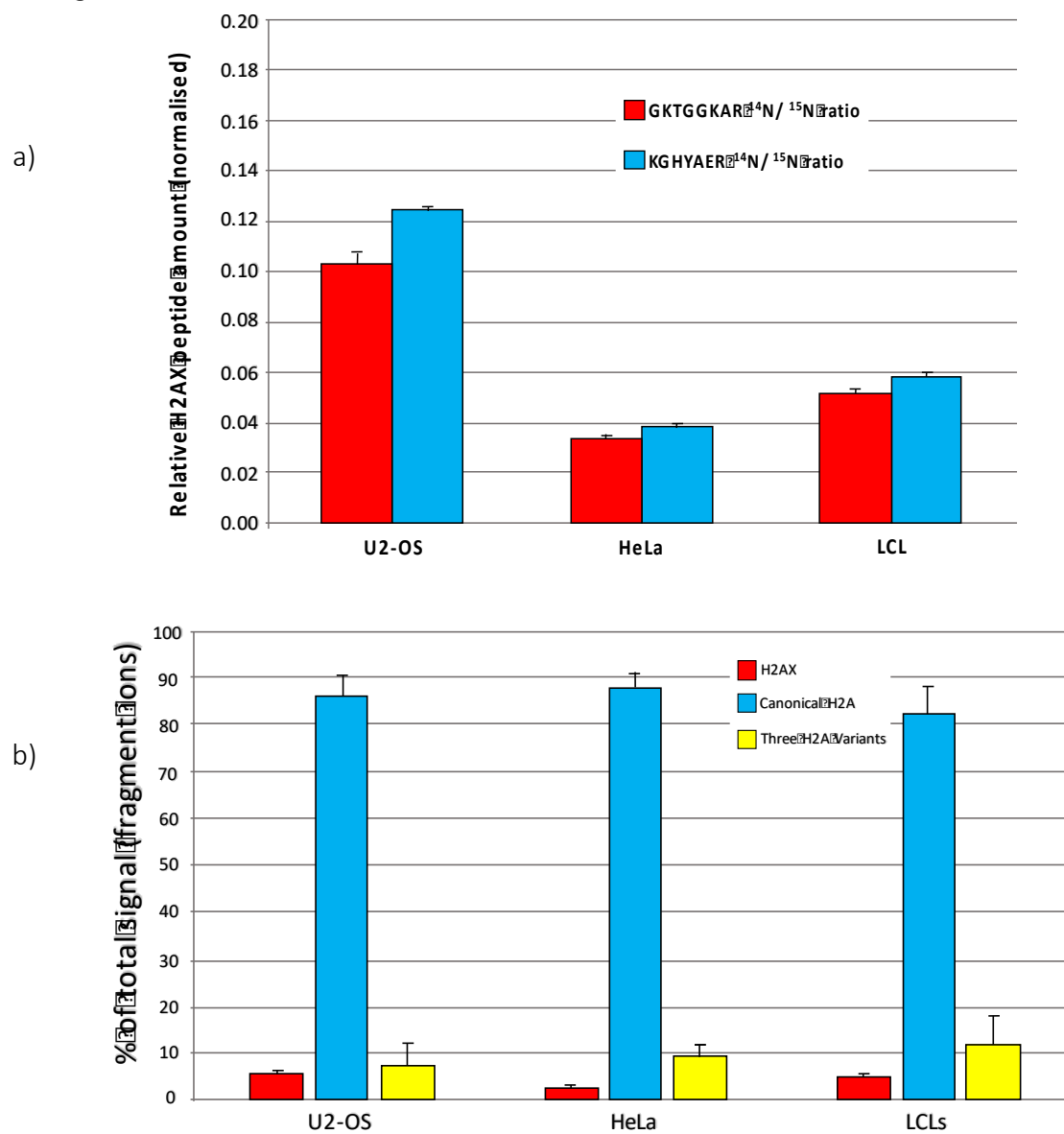
- a) Endogenous S139 phosphopeptide was enriched from undamaged U2-OS, relative to the unmodified peptide. Ions fragmented from singly charged precursors were used for quantification.
- b) ATQASQEY and ATQAS[p]QEY synthetic peptides were successfully enriched from 100x excess BSA peptides.
- c) H2AX C-terminal peptides
- d) Precursor ions from b) shown with relative intensity pre- and post-enrichment.
- e) ATQASQEY[p] synthetic peptide enriched successfully after being spiked into U2-OS acid extract. 50 µg acid extract sample prepared and 0.1% of it was analysed pre-enrichment. (LHS) Using pseudo-SRM both ATQAS[p]QEY and ATQASQEY[p] were detectable pre-enrichment. Precursor ions were not detectable (Data not shown). (RHS) Both ATQAS[p]QEY and ATQASQEY[p] synthetic peptides are identified with much improved signal-to-noise ratio post-enrichment.

3.2.3. H2AX abundance

The relative abundance of two peptides unique to H2AX (GKTGGKAR and KGHYAER) were compared with those of similar peptides in other H2A isoforms. Three cell lines were used for the assay: U2-OS, HeLa, and LCLs. The peptides were propionylated before and after trypsinization, so as to prevent trypsin digesting at lysine residues. A known amount of heavy labelled ¹⁵N recombinant H2AX and H2A1B (Methods and Materials) were spiked into the samples to account for variation in sample preparation and analysed in technical triplicate. **Figure 3.4a** shows that across the cell lines, U2-OS contains the larger of relative abundance of H2AX: Threefold and twofold more compared to HeLa and LCLs respectively. The endogenous unique H2AX peptides were normalised to the spiked-in ¹⁵N recombinant H2AX across all the cell line. The relative abundance was consistent with the two H2AX unique peptides across the cell lines.

The two H2AX unique peptides (KGHYAER/GKTGGKAR) were then used to quantify the levels of H2AX in the proteome compared to other H2A variants. Three peptide sequences were used for this analysis: GKTGGKAR/KGHYAER (H2AX), GKQGGKAR/KGNYAER (canonical), and GKGGVKRKGNYSER (one/three variants). The peptides were concatenated to correct for ionisation efficiencies. **Figure 3.4b and Table 3.1** show that U2-OS possess the largest percentages of H2AX, then LCLs and HeLa cells having the least amount of the cell lines analysed. Even though the percentages were not the same for both peptides, they are consistent in demonstrating H2AX levels across the three cell lines. Four H2A variants were not analysed as they did not possess a specific sequence to compare against.

Figure 3.4. H2AX Abundance



3.4 H2AX Abundance

- a) H2AX relative amounts were measured across three cell lines. U2-OS, HeLa and LCL extracts were spiked with ¹⁵N recombinant H2AX and H2A1B. Diagnostic peptide ratios were quantified. To control for discrepancies in prepared histone starting material, H2A1B GKQGGKAR was used to normalise the diagnostic peptides across samples.
- b) Percentage of H2AX variants compared to other H2A isoforms. U2-OS, HeLa and LCLs H2AX levels were measured relative to other H2A variants. The levels were corrected for ionization efficiency. Fragments from doubly charged precursors were used for quantification. N = 2. Technical triplicate were used to analyse each of the two biological replicates.

Table 3.1. H2AX Abundance

Measured peptide	Protein(s)	U2OS	HeLa	Lymphoblast
GKTGGKAR	H2AX	3.1	1.0	1.9
GKQGGKAR	11 H2A (1A, 1B, 1C, 1H, 1J, 1D, 1, 2A, 2C, 3, 2B)	96.3	97.7	97.6
GKQGGKVR	H2AJ	0.6	1.3	0.5
Not measured	H2AB1, H2AB2, H2AV, H2AZ	n.m.	n.m.	n.m.
Total	17			
KGHYAER	H2AX	6.4	2.6	5.1
KGNYAER	9 H2A (1A, 1C, 1H, 1J, 1, 2A, 2C, 2B, J)	91.1	88.4	82.9
KGNYSER	H2A 1B, 1D, 3	2.4	9.0	11.9
Not measured	H2AB1, H2AB2, H2AV, H2AZ	n.m.	n.m.	n.m.
Total	17			
	H2AX average	4.7	1.8	3.5

Table 3.1. H2AX Abundance (Legend)

Table of relative amounts of various H2A isoforms. Detection of GKTGGKAR was 1.1x better than GKQGGKAR the canonical H2A sequence. KGHYAER, peptide sequence unique to H2AX, was detected with 1.7x stronger signal, compared to the canonical KGNYAER. The relative percentages have been corrected to take into account different ionization efficiencies. N = 2. Technical triplicates were used to analyse each of the two biological replicates. Values for standard deviation are in brackets.

n.m. = *not measured*

3.3. Discussion

H2AX C-terminal phosphorylation levels were able to be quantified using in-house developed pseudo-SRM assays. Due to two structural reasons, phosphorylation of both S139 and Y142 are challenging modifications to analyse. Normally, the last amino acid after a tryptic digestion is basic in nature: arginine or lysine. The basic residues would assist in balancing the charges brought on by modifications to the peptide -OH side chain. Each additional phosphorylation to the peptide will increase the negativity by -1. Together, the lack of a basic molecule and receiving phosphorylation has an additive effect of decreasing the ionisation efficiency in positive mode mass spectrometric quantification assays, resulting in the peptide signal more difficult to detect. Nevertheless, both phosphorylated Serine 139 and Tyrosine 142 were able to be detected in both singly and doubly charged species, which can be fragmented into fragment ions that make it easier for quantification. Even though the two phosphorylated peptides are isobaric in nature, they do not co-elute from the nano-LC, therefore, there is no overlap in the diagnostic fragment ions from both species which would make quantification of both modifications more nebulous.

We see discrepancies between both U2-OS and LCL in γ H2AX levels after being irradiated with 20 Gy of ionising radiation. The large differences between the two cell lines in γ H2AX levels has two possible explanations: sample prep differences and biological differences. The histones from LCLs were obtained with a Micrococcal nuclease digestion of isolated nuclei, and the histones from U2-OS were obtained with an acid extract of the two nuclei. Points that would need to be taken into consideration before making a judgment in stating that there are biological differences between LCLs and U2-OS in regards to

phosphorylation S139 levels: the samples were prepared at different times, and two different methods of histone preparation were used. While an acid extract, in theory, should extract all the histones from the chromatin as it is a crude method of extraction, it is difficult to say where the histones would be coming from in a Mnase digestion. Cowell et al. demonstrated previously that H2AX S139 is more likely to be phosphorylated in much more open euchromatic regions of the genome (Cowell et al. 2007). Using MCF7 cells, they observed that at the 30-minute time point two-thirds of γ H2AX foci do not overlap with HP1 α . However, more recent studies suggest that HP1 α is recruited to sites of damage (Baldeyron et al. 2011). Nonetheless, Mnase does not digest and release nucleosomes uniformly across the genome, especially when compared to acid extraction that is not so discriminating in where nucleosomes are obtained from. What is consistent in both γ H2AX assays with LCLs and U2-OS is that there is a peak of γ H2AX signal around 30-minutes after damage. In both irradiated cells, there is a ten-fold increase in γ H2AX compared to cells that were not subjected to IR. In U2-OS cells, phosphorylated Y142 H2AX was not detected in undamaged or damaged cells, even after enriching for phosphorylated peptides from a copious amount of starting material, 50 μ g. We were able to use the same method to enrich for synthetic phosphorylated Y142 peptides with good signal. It is possible that global Y142 phosphorylation levels are below 1% and below the levels of detectability with the instrument used in the analysis. Phosphorylated S139 was enriched with a signal to noise ratio of 100:(1e6/9000). Furthermore, we identified phosphorylated S139 has 20-40x lower abundance natively in the genome (in relation to damaged chromatin), indicating that the lowest threshold for detection of γ H2AX is between 0.125-0.25%. When taking ionisation efficiencies correction into account; phosphorylated Y142 peptide is detected 3.2x less than the phosphorylated S139

peptide, then that leaves an upper ceiling detectability of 0.8% phosphorylation Y142 at a genome level. Other proteomic studies that looked into phosphorylated Y142 also had difficulty with detecting the modification. Both in the presence of damage where S139 phosphorylation was also detected and in the absence of DNA damage (Olsen, Vermeulen et al. 2010; Rigbolt, Prokhorova et al. 2011). Therefore, it is not farfetched to say that the presence of phosphorylated Y142 is so low that it is below the threshold of detectability with mass spectrometry. An estimate that takes into account ionisation efficiencies of the phosphorylated species would suggest that phosphorylated Y142 was at least below 1% of H2AX in undamaged U2-OS. Even though U2-OS cells were used in this case, other cell lines were used, and it is possible that those cell lines, such as HEK 293T, mouse 3T3 and mouse embryonic fibroblasts, express the WSTF kinase more efficiently or the WSTF kinase is more potent in those cell lines. However, with Y142 phosphorylation being implicated in both apoptosis and the DDR, it is not likely there would be large variance across different cell lines or within highly conserved mammalian systems.

Of course, the implication of this finding is that it questions the findings of studies that used immunoblotting systems to quantify or detect new functions of Y142 phosphorylation. Xiao et al. raised an antibody that binds to Y142 phosphorylation as an epitope site. However with there being a ratio of >100:1 mod:pY142, a small margin of error could result in potentially erroneous results with a likelihood of false positives being given in data outputs (Xiao, Li et al. 2009). The low abundance of phosphorylated Y142 also affects the paradigm in which Y142 was previously thought to be involved with both the DDR and apoptosis. With phosphorylated Y142 being present at the genome at below 1%, means that for both S139 and Y142 to be phosphorylated simultaneously on the same H2AX molecule would be an even lower percentage (Cook, Ju et al. 2009; Singh,

Basnet et al. 2012). The only time it would seem to be conceivable for both Y142 and S139 to be phosphorylated would be in an environment that possesses large amounts of DNA damage, when γ H2AX is to be the majority (Brown, Eykelenboom et al. 2012). Using site-directed mutagenesis to elucidate further functions of Y142 is complicated by the fact that when Y142 is mutated, Binding of MDC1 to γ H2AX is perturbed (Xie, Hartlerode et al. 2007). Even though WSTF is dispensable during meiosis and is not required for male fertility (Broering et al. 2015), removal of WSTF reduces γ H2AX formation and hinders the DDR. It remains to be determined if the DDR depends on phosphorylated Y142 in some form or another or whether the chromatin remodelling role of WSTF is required for the DDR (Xiao, Li et al. 2009). The findings here indicate that WSTF has an indirect effect on the DDR.

H2AX was also found to be present at around 2-5% in three different cell lines. Out of 24 H2A variants, only 17 were analysed, which comprises the vast majority of H2A variants, the percentage of H2AX will be slightly inflated as a result. However, it is not orders of magnitudes different to the real number and the level calculated here was close to the correct figure. U2-OS was found to possess the most amount of H2AX compared to LCL and HeLa. To suggest that the increased response to 20 Gy of IR is due to an increasing number of H2AX would be premature as consistent controls are required to be carried out first.

1.1. Chapter 4: The Development and Optimisation of γ H2AX-IP-MS Method

4.1. Introduction

In Chapter 3, the only changes in HPTM levels that could be detected via LC-MS upon treatment of cells with ionizing radiation was γ H2AX. As a result, it became apparent that a new method of enriching histones from damaged chromatin would be required. The hypothesis was that the main issue with using a method solely based on an acid extraction for a DNA damage proteomic screen was that the acid extraction does not filter out non-DNA damaged chromatin, resulting in excessive background noise from histones that were not near the site of damage (**Figure 4.1a**). Apart from HPTM changes in γ H2AX, other HPTMs did not show any changes as a result of damage (Chapter 3). It was further hypothesised that histones from the rest of the genome were masking the actual signal of DDR epigenetic changes. Therefore, a system that would only target DDR histones was required. Several approaches were considered, but a two-pronged strategy was settled to reach this goal. The first approach was to pull-down different DDR proteins that are known to bind to damaged chromatin, such as RAD51 and 53BP1, and examine the PTMs of associated histones (**Figure 4.1b**). The second approach was to pull-down histones associated with damaged chromatin using damage-specific antibodies (**Figure 4.1c**). To identify the best pulldown method prior to LC-MS, both native pull-down and tagging a protein were investigated.

The abundance of histones within the genome was always going to be a double-edged sword: they would be relatively more straightforward to enrich compared to other proteins, but their copious amounts would also result in increased susceptibility in obtaining misleading and/or undetectable data. Therefore, it was imperative for a high quality proteomic screen to develop a method that had the right balance of enrichment combined with a good signal-to-noise ratio. Other factors would also come into play and how they were negotiated will be elaborated further on. The motivation to develop a method that would give a better proteomic read out of the DDR-related HPTMs was because the crude method of preparing samples for mass spectrometry with just an acid extraction from isolated nuclei resulted in a reduced signal-to-noise ratio for detecting changes in HPTMs apart from γ H2AX.

4.1.1. Rationale of Choice of DDR Proteins

There were two approaches to consider when enriching for damaged chromatin. Either a protein that is known to associate with chromatin upon DNA damage could be targeted, or the other option would be to enrich for damaged chromatin directly (**Figure 4.1b and 4.1c**). The initial intention was to develop both in tandem to hedge our bets in case one approach was not successful.

Only two DDR proteins came to mind when considering which one to use for enriching: 53BP1 and RAD51. Initially one would think to use one of the kinases responsible for phosphorylating γ H2AX. However, ATM, ATR and DNA PKcs have approximate molecular weights of 350 kDa, 317 kDa, and 469 kDa, respectively (**The UniProt Consortium 2017**). Even though the kinases of γ H2AX would be at the site of damage immediately, their

relatively large protein sizes would also make them more challenging to enrich with a decent yield for a proteomic screen.

Figure 4.1. Amplifying DDR-HPTM For Proteomic Screen

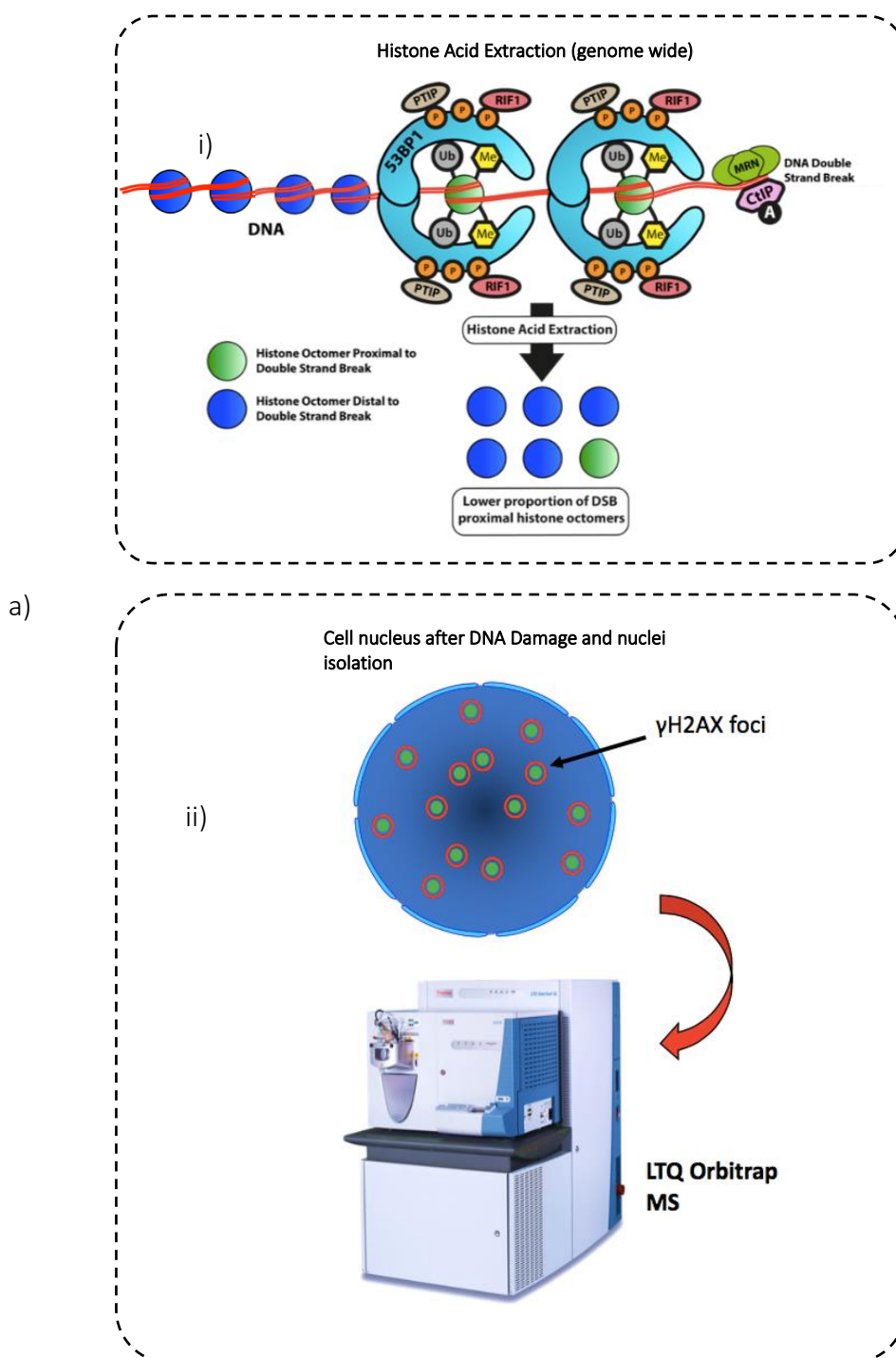


Figure 4.1 Legend Introduction

- a) (i) Preparing histones via acid extraction does not allow to for the distinction of nucleosomes that are close to the site of damage and those that are far from the site of damage. Quantifying HPTM's close to the site of the damage becomes a problem as a result. The indiscriminate histone extraction is not a hindrance when a genome wide assay is desired.
- (ii) Due to the chemical nature of acid extractions, all of the histones within the genome of a cell will be pulled down together regardless of where they are located in the nucleus. H2AX occupies approximately 5% of H2A variants, and the percentage of these containing the phosphoS139 modification would be even fewer. Therefore, the MS/MS signal obtained from nucleosomes in a DNA damage environment will be overwhelmed by the >90% of nucleosomes not containing γ H2AX, making quantification of HPTM close to or at the site of DNA damage unfeasible. DSB, double strand break.

4.1. Introduction (continued)

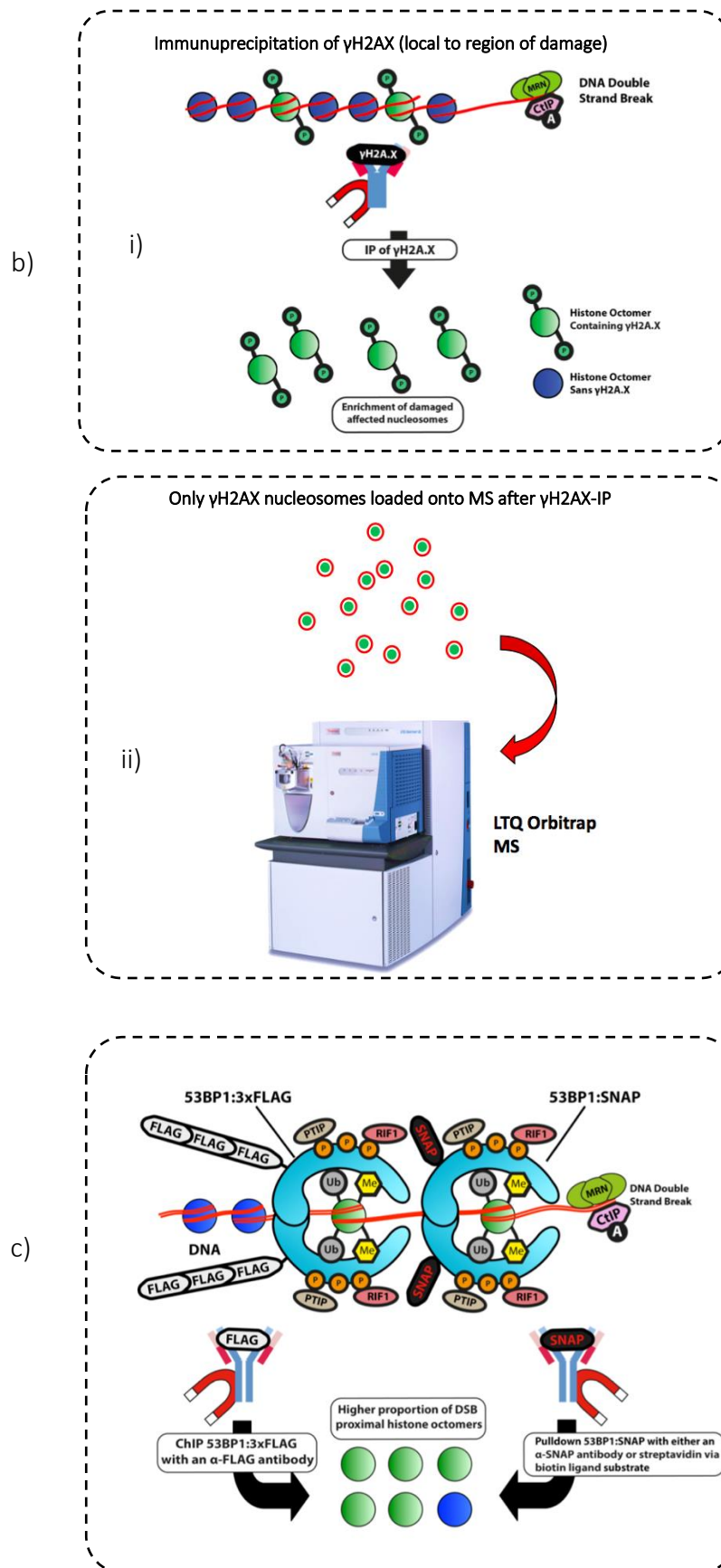


Figure 4.1 Legend Introduction (continued)

- b) (i) Using an α - γ H2AX antibody, you can enrich for nucleosomes that are close to the site of damage, or at the site of damage exclusively. The DNA bound nucleosomes together can be digested using micrococcal nuclease, thus releasing the nucleosomes.

(ii) The MS/MS signal obtained from running histones enriched from sites of damage exclusively will give better quantification in regards to HPTMs that are damage related.
- c) 53BP1 is known for binding to ubiquitinated histone H2AK13/K15 and histone H4K20me1/2. 53BP1 foci are synonymous with double strand breaks. Thus, it should be possible to enrich for nucleosomes containing DSB-associated histones by isolating 53BP1 after damage has occurred with either native immunoprecipitation or with the aid of crosslinking agents. By enriching damaged associated nucleosomes, the signal for damaged related HPTMs should be enhanced compared to MS/MS signal with no enrichment.

The other two DDR proteins that could be used are RAD51 and 53BP1 which have a molecular weight of 37 kDa and 214 kDa, respectively **(The UniProt Consortium 2017)**. If it were a just a matter of protein size, RAD51 would have been chosen, but there are also time-sensitive molecular biological logistics that need to be taken into account. The chromatin modification that takes place as a response to DNA damage is an ever-changing series of events, so determining what is taking place at a temporal scale was preferred **(Cao et al. 2016)**.

For this reason, RAD51 is ruled out as a protein of choice for pulling down damaged chromatin, as RAD51 comes into play later in the DDR compared to 53BP1 and time-sensitive events would be missed as a result **(Goodarzi et al. 2008)**. Not only would early events be missed but RAD51 efficacy as a pull-down candidate starts once homologous recombination has been chosen as the repair pathway, which mainly occurs in S-phase and G2-phase. Most of the vegetative cells remain in G1, about 80%. Even if a double strand break takes place in G2, 80% of the time, the repair pathway is chosen is still NHEJ **(Kakaroukas & Jeggo 2014)**. Therefore, 53BP1 is the most suitable choice to pull down damaged chromatin that is both nascent and proximal. 53BP1 is an early response protein, and its response is quick with many molecules recruited to the site of damage that it used as a readout for double-strand breaks in immunofluorescence **(Panier & Boulton, 2014)**. There is also the added advantage that 53BP1 binds to several histone PTMs: H2AK13ub/K15ub, H4K20me2, H3K79me2, and γ H2AX **(Panier & Boulton, 2014)** **(Baldock et al. 2015)** **(Wakeman et al. 2012)**. Also from a technical point of view, the three isoforms of 53BP1 are all similar in size with only a difference of 0.5 kDa between them. In contrast to RAD51, which has four different isoforms ranging between 25 kDa

and 37 kDa in size (**The UniProt Consortium 2017**). The issue would be that it would make interpreting western blots and Coomassie stained SDS-PAGE gels more difficult if multiple bands are corresponding to the same protein. Unlike γ H2AX, which is present at both single-strand breaks and double strand breaks, 53BP1 only binds to double-strand breaks. A potential issue with 53BP1 is that it does not persist long enough for late repairing breaks so it might not be sufficient for studying HPTM changes in late repairing breaks.

The other approach was to directly enrich for nucleosomes from damaged DNA with the use of an antibody that has a specific PTM as its epitope. Compared to pulling down a DDR associated protein, this method should be less of an arduous task as there would be no need to modify the native histones. Three HPTMs are indicative of DNA damage: H2AK13ubi/K15ubi, H4K20me2 and γ H2AX (**Cao et al. 2016**). The three modifications are covered much more in-depth ex-supra but in-brief: γ H2AX is much more characterised, and its formation after the damage is so rapid compared to ubiquitination of H2AK13/K15 and dimethylation of H4K20. Therefore an IP of γ H2AX would yield a higher enrichment resulting in better signal to noise ratio. Another factor is what biochemical reagent to use. Due to γ H2AX being used an identifier of damaged chromatin in many different assays there is a plethora of antibodies from many different manufacturers to choose from which will increase the chances of using an antibody suited for the task. Antibodies for ubiquitinated H2AK13/K15 are less common among the available assortment of commercial antibodies. Antibodies specific for ubiquitin are used to determine if the protein of interest has been ubiquitinated compared to an antibody for γ H2AX which is specific for the PTM and the protein of interest. A single antibody is available from Merckmillipore (catalogue number: EDL H2AK15-4) that can apparently recognise

ubiquitination at H2AK15 (no mention of H2AK13) at the time of writing. However, it has not been tested for immunoprecipitation applications. Another consideration is that there are so many variants of histone H2A. It is not certain whether an antibody would be recognising all of the ubiquitinated H2AK13/K15 as an amino acid difference or two between two or more variants could result in the antibody only recognising some variants of histone H2A while others would be ubiquitinated but also missed.

Histone H4 only has one variant which is advantageous. Also, a large number of antibodies exist for H4K20me2. However, the problem with using H4K20me2 to enrich for damaged chromatin is that the vast majority of H4K20 is dimethylated outside of damage (Pesavento et al. 2007). Despite several studies showing that H4K20 dimethylation increases following damage, there is high proportion of background H4K20me2 not originating from damage sites, and avoiding the enrichment of background non-damaged nucleosomes is the purpose of this exercise.

When taking into account the size of a mammalian genome and the number of nucleosomes it encompasses and that only a few of those nucleosomes would exhibit changes as a result of DNA damage, the more apparent option would be to pull down γ H2AX. It also cannot be understated the usefulness of γ H2AX possessing a unique tail (Chapter 3) that no other variant of H2A has in their amino acid sequence; this rapidly decreases the chances of having a false positive from a α - γ H2AX antibody. Therefore, the decision left was to use γ H2AX for: its quick response following damage; its persistence at the late-repairing break; the fact that it is a highly specific and unique PTM; and the vast abundance of readily available commercial antibodies to choose from for developing the method.

4.1.2. Potential Technical Issues

It should be mentioned that by introducing a biochemical affinity enrichment step to the proteomic method, not only was another potential point of weakness being added whereby things could go wrong, but the mechanical precision of the mass spectrometer was being sullied with the mercurial temperament of biological reagents such as antibodies. Essentially the purification step is the most vulnerable part as that is where most non-specific contamination can take place. Antibodies are notorious for their issues which can be classified into one of three categories: non-specificity, batch-to-batch inconsistencies and incorrect application (**Figure 4.2**). Antibodies could recognise unintended epitope sites that are similar in either both sequence and modification relative to the target protein/site. Immuno-assays of histone post-translational modifications could be problematic, especially with the issue of cross-reactivity as similar PTMs are dotted around across various histones. For example, an antibody for H3K9me3 or H3K9ac could potentially recognise H3K4me3 and H3K14ac respectively, which are lysine residues that very close to each other. Resulting in false positives leading to erroneous findings (**Figure 4.2d**).

Tjeertes et al. acknowledge different results from their lab group compared to another separate lab group concerning whether or not H3K56ac increases after DNA damage (**Tjeertes et al. 2009**). They explain that the discrepancy is due to the antibodies targeting the same epitope used by the different labs coming from two different manufacturers. Numerous other studies have questioned the results from antibodies. Egelhofer et al. (2010) (**Egelhofer et al. 2010**) tested nearly 250 epigenetic antibodies that had been

utilised in previous studies and found that approximately a quarter were not specific and would bind to multiple sites. Historically, antibodies have also been undependable in immunohistochemistry. Buchalow et al. (2011), demonstrated that excessive blocking did not prevent the antibodies from binding to non-specific sites (**Buchwalow et al. 2011**). In a separate study, over 20,000 commercial antibodies were analysed, fewer than half were found to have a minimum threshold of reliability to observe the distribution of proteins in slices of tissue (**Berglund et al. 2008**).

Figure 4.2 Common problems with antibodies

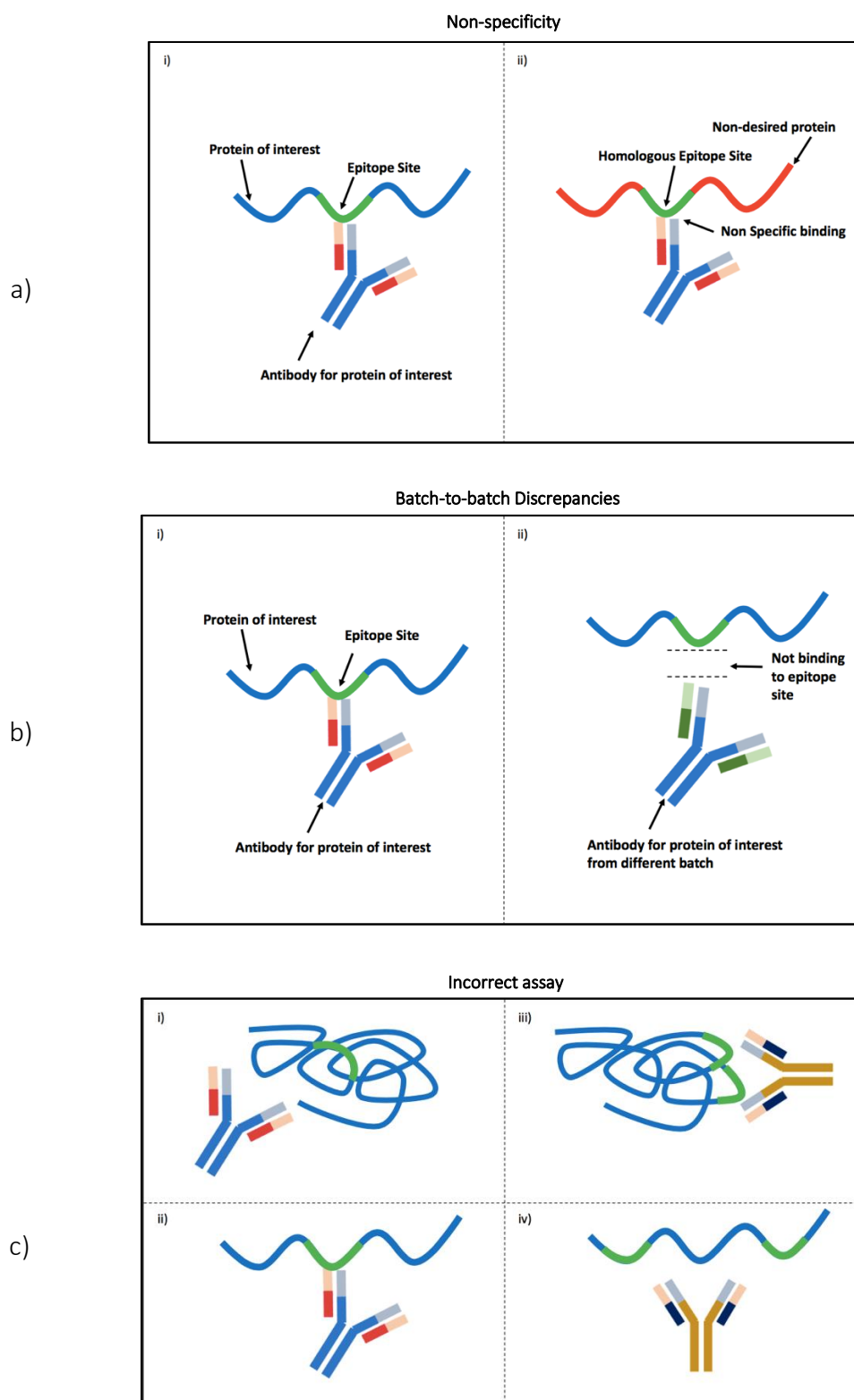
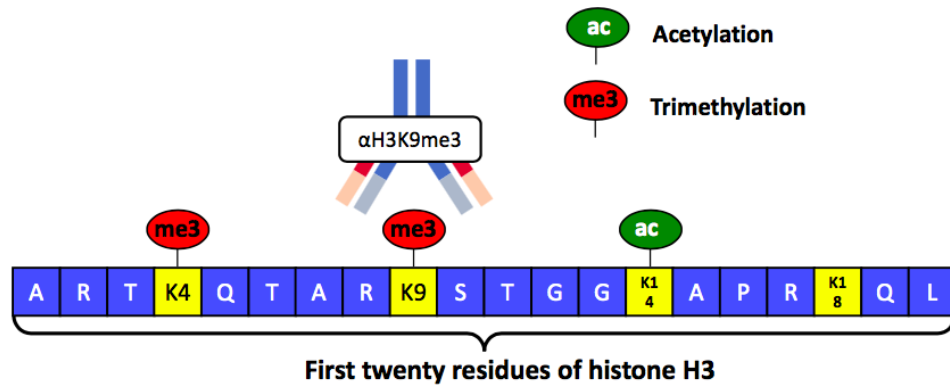


Figure 4.2 Legend: Common problems with antibodies

- a) Non-specificity: an antibody is deployed that was intended for a protein of interest but a non-desired protein is also recognised by the antibody because it shares sequence homology with the original protein that corresponds to the epitope site of the antibody. To avoid such issues, you would test the antibody with stringent controls such as knockout cell lines to ensure the antibody only has one antigen.
- b) Manufacturing Inconsistencies: Different lot numbers for the same product (identical catalogue numbers) from a single manufacturer could provide inconsistent results from one experiment to the next: the antibody worked optimally in recognising an antibody **(i)**. When the lab stock was depleted and new stock was procured, it results in a different result for the assay **(ii)**. To circumnavigate this potential pitiful, you could keep to using only monoclonal antibodies if possible for your protein of interest. Otherwise conduct an audit with the supplier for their quality assurance if such discrepancy is noticed.
- c) Incorrect assay: Depending on the antibody, they could recognise the same peptide in its denatured form exclusively because the epitope site is buried with in the core of the folded protein **(i)** but when the peptide is denatured it exposed for the antibody to recognise **(ii)**. You would use an antibody that functions in this manner in applications where the protein ends up being denatured such as in Western Blotting. An antibody could also recognise a given protein in its native form where the epitope site is formed by two domains being brought closer due to the proteins 3D conformation resulting in a folded-only epitope **(iii)**. However, when the peptide is denatured, the antibody no longer recognises the protein because the epitope is no longer present **(iv)**. Antibodies could be useful in this scenario for applications that use native proteins such as immunohistochemistry.

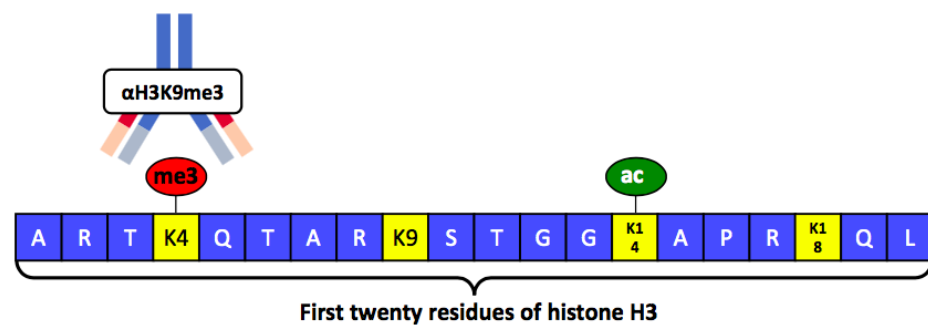
Figure 4.2 Common problems with antibodies (continued)

Scenario of antibody not functioning as intended resulting false data



Antibodies can be non-specific; recognising wrong epitope site.

d)



H3K9me3 antibody recognises the wrong site, thus giving a false positive.

Figure 4.2 Legend Common problems with antibodies (Continued)

- d) Antibody non-specificity. (LHS) An antibody raised against H3K9me3 will detect the correct epitope site when the trimethylation modification is present on the residue. (RHS) Potentially, an antibody can bind to a similar epitope site that possesses the same residue at a different site with the identical modification. In this case, the α -H3K9me3 antibody gives a false positive by recognising the trimethylation on the nearby H3K4 residue. Antibodies could also recognise methylation at a specific site without distinguishing how many methyl groups were deposited on the amino acid residue. For example, an antibody binding to both H3K9me2 and H3K9me3.

In another case of non-specificity of antibodies giving inconclusive results, Michel et al. found that the majority of signalling receptor antibodies they tested recognised multiple proteins thus they did not give any confidence in being able to differentiate one receptor from another (Michel et al. 2009). It is one thing to cause hindrance to a scientific study, but when there is a translational application intended, it can be disheartening to find out that close to 90% of separate cancer studies were not reproducible as a result of antibodies (Begley & Ellis 2012).

What needs to be taken into consideration as well is that the most common uses of antibodies, i.e. western blotting and immunofluorescence, do not give direct readouts: it is assumed that the secondary antibody has bound to a primary which itself has bound to a given protein of interest. An issue more so for polyclonal antibodies but still possible with monoclonal antibodies are discrepancies in efficacies from one batch to the next could be introduced at the manufacturing stage, when a given antibody is raised within a new stock of animals or even bled on a different day. The problem arises when different batches of an antibody are advertised as suitable for specific applications on the manufacturer's website and given the same catalogue numbers.

Antibodies are not standard chemical reagents manufactured on an assembly line. They are biological reagents that are more susceptible to variance due to their nature. One needs to maintain an air of scepticism when procuring antibodies from any given manufacturer. Prassas et al. lost nearly \$500,000 and two years chasing a phantom result due to a mischaracterisation of a commercial antibody-based protein detection kit for CUZD1, a protein found to be overexpressed ovarian cancer. They found that the kit was

detecting the wrong protein altogether: CA125 (Prassas & Diamandis 2014). To ensure that this experiment did not fall into the same trap, it was decided to implement a stringent design of experiment (DOE) where a series of quality control tests were carried on multiple antibodies from different manufacturers (Bordeaux et al. 2010).

Another potential problem to consider is whether the antibody is suitable for the application that is being used. As mentioned, antibodies are utilised in assays that are as varied as immuno-precipitation to immunofluorescence. For example, just because the manufacturer states the antibody is specific for immunoprecipitating γ H2AX, it does not mean the antibody would be suitable for detecting γ H2AX in an immunofluorescence application. Some antibodies have been designed to identify epitope sites when the protein is in its native structure which would make the antibody suitable for immunofluorescence. The epitope, in this case, could be two sequences that are within proximal distance when the protein is folded. However, for applications that require the antibody to be denatured such as for western blotting, the antibody as mentioned above would not be suitable as the sequences would be too far apart to be recognised as an antigen. Therefore, an antibody that is more suitable for unfolded polypeptides would be more suitable (Figure 4.2a).

The frequency of antibody validation has increased in recent years. Antibody validation is the process of reporting on which applications the antibody is suitable for and whether there is cross-reactivity, state whether the antibody has been recalled, etc. There is increased collaboration from academia, industry and the biotech manufacturers in ensuring that antibodies function as described. However, the onus of assessing the

suitability of the antibody remains on the investigator to ensure that the antibody is applicable. In this project, the most vulnerable part would be the immunoprecipitation step of the IP-MS γ H2AX/53BP1-3xFLAG. Even though trypsin and Mnase are also biological reagents, their function is relatively more binary: they either cut or do not cut at specific sites of protein and DNA, respectively. Therefore, to determine if they are functioning or not would require an assay with a rudimentary readout, i.e. running an agarose or SDS-PAGE gel.

4.2. Results

4.2.1. 53BP1-3xFLAG Pull Down

Transient 53BP1-3xFLAG and 53BP1: SNAP cell lines were established. Stable cell lines of 53BP1-3xFLAG were also established, but stable 53BP1:SNAP cell lines could not be established as the cells failed to survive the neomycin selection. Expression of 53BP1-3xFLAG was confirmed by western blot and immunofluorescence experiments (**Figure 4.3a, 4.3b**). Immunofluorescence indicated that the native function of 53BP1 was not compromised when tagged with FLAG as α -FLAG and α - γ H2AX exhibited overlapping foci (**Figure 4.3b**). 53BP1-3xFLAG was successfully able to be enriched with an IP using an α -FLAG antibody (**Figure 4.3d**). However, putative 53BP1 bands were barely visible after Coomassie staining, suggesting that the quantity of IP'd protein was quite low. Nevertheless, bands at approximately 190-245 kDa were excised and prepared for mass spectrometry analysis (**Figure 4.3e**). Bands from the 53BP1-3xFLAG cell line showed 1-2% peptide coverage of 53BP1. In the untransfected cell line, 53BP1 was not detected at all via mass spectrometry (**Figure 4.3f**).

The streamlined option to enriching for 53BP1 would be to procure an antibody specific for 53BP1. If an unmodified protein was purified, it would save time in setting up stable cell lines. It would also add more confidence in the protein being pulled down. Another potential advantage of using antibodies for the protein themselves is that polyclonal antibodies could be used, which would have multiple epitope sites so instead of having one antibody per protein molecule one could have multiple antibodies per protein molecule, which would make it easier to pull down larger proteins such 53BP1. The caveat of using polyclonal antibodies would be that they could lead to non-specific binding.

However, good quality monoclonal antibodies specific for 53BP1 can cost 3-4 times as much, if not more, as an antibody raised for a common tag. At the time of writing, Abcam sells α -FLAG (catalogue number: ab49763) which costs £285 for 100 μ L at a concentration of 1 mg/mL, while the highest rated α -53BP1 antibody from Abcam (catalogue number: ab175933) costs £289 for 100 μ L 0.286 mg/mL. Sigma Aldrich sells an α -FLAG antibody with quantities of 0.2 mg for £332 or even 1 mg for £752, which substantially reduces the costs when purchased in bulk. The α -FLAG antibody had already been well characterised for ChIP, by both our lab and others in previous studies. Testing a suitable α -53BP1 for IP, would not only be much costlier as multiple α -53BP1 antibodies from multiple manufacturers would require assessing, a lot of time would be spent trying to optimise the condition for the untested antibody. Therefore, it was neither economically feasible or productive use of time to use an antibody specific for 53BP1. Even though it may seem both gauche and meta to discuss the costs of antibodies in a scientific study, one still needs to consider that when developing a novel method the consequences of excluding others from attempting to reproduce the method for their application solely based on financial ceilings.

However, there are also some potential issues with tagging a protein: the original function of the protein being altered or even the tag itself being inaccessible due to how the protein folds. If the wrong tag was also selected, e.g. HIS tag, it could result in non-specific proteins being pulled down alongside the tagged-protein due to some cellular proteins containing multiple histidines adjacent to each other in their amino-acid sequence (Bornhorst & Falke, 2000). The lab already possessed a recombinant plasmid containing 53BP1. Attaching two different tags to 53BP1 via cloning was chosen and to transfect the tagged version into U2-OS cells, with the intention of creating a stable cell

line. The cloning was successful and was confirmed by sequencing. The plasmid maps detail where the tags were cloned in (**Figure 4.2c**). When the U2-OS cell lines were transfected, both tags could be expressed transiently (**Figure 4.2a**). For whatever reason, the SNAP-tagged version of the cell line could not survive long term neomycin selection. The lack of a SNAP-tagged cell line was not such an issue as by using two different tags, the chances of using a compatible tag with 53BP1 were increased. From then on the focus was on the FLAG tag version of 53BP1. The initial concern of having the tag alter the original function of 53BP1 was ablated as the FLAG tag was observed to have overlapping foci with γ H2AX in etoposide challenged U2-OS cells that had been transiently transfected with the 53BP1-3xFLAG plasmid (**Figure 4.3b**).

Figure 4.3. 53BP1: 3xFLAG Pull Down

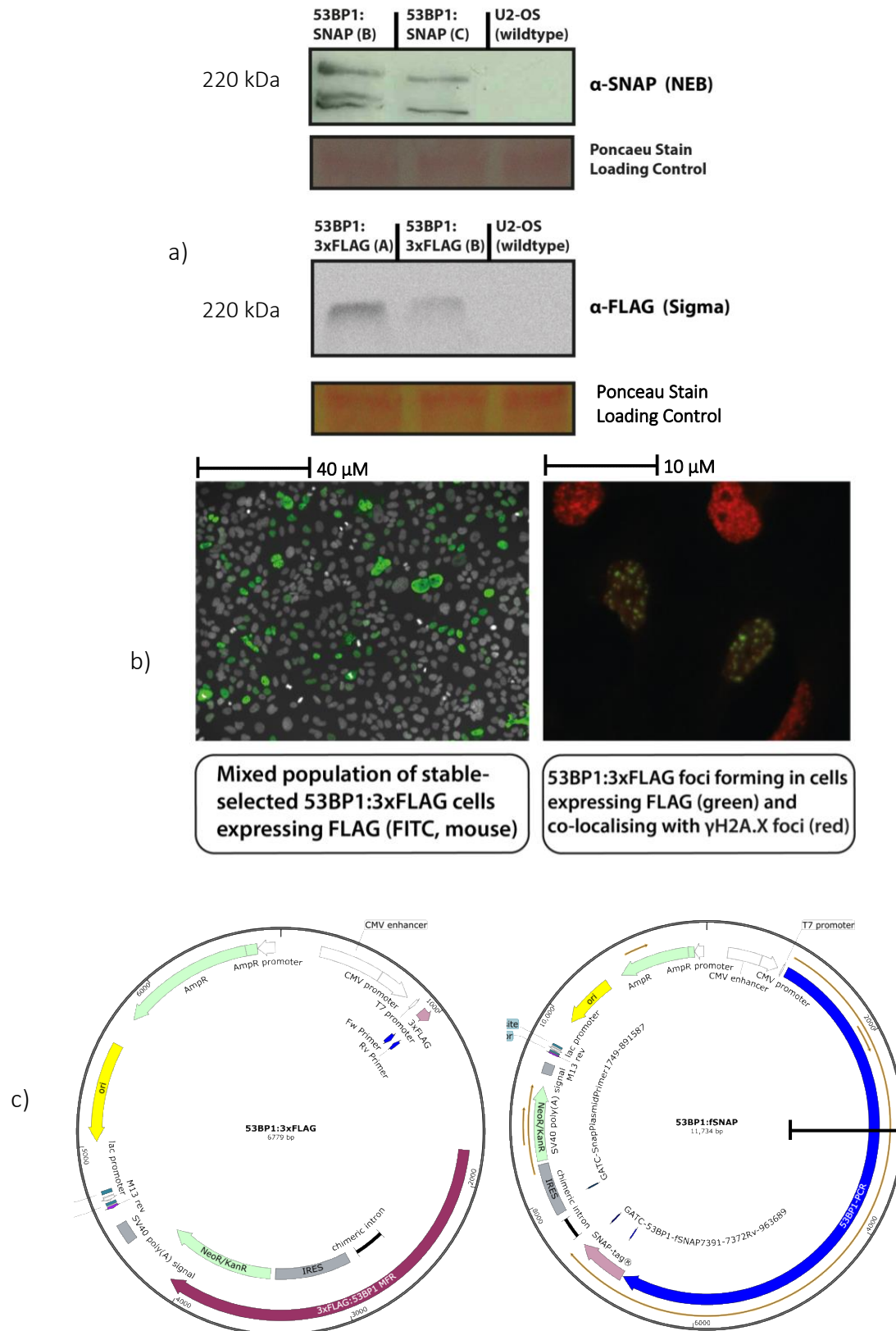


Figure 4.3: 53BP1: 3xFLAG Pull Down

- a) Western blot of transient cells expressing 53BP1:SNAP (top) and 53BP1:3xFLAG (bottom). SNAP and FLAG tag were both probed to determine whether cells could express the tagged version of 53BP1. Ponceau staining used as a loading control. Multiple SNAP due to degradation. U2-OS wild type were not transfected.
- b) Stably selected mixed population of U2-OS cells expressing 53BP1:3xFLAG (left). Confirmation of FLAG DSB foci formation, suggesting that the DNA damage-induced chromosomal localisation of 53BP1 to DSBs was not compromised. Cells were treated with 20 μ M etoposide for 1 hour. Microscopy was performed on the ScanR microscope.
- c) (Left) 53BP1:3xFLAG plasmid used to transfect U2-OS cells. 3xFLAG tag placed at the C-terminus of 53BP1. (Right) 53BP1:fSNAP plasmid used to transfect U2-OS cells. SNAP tag placed at the C-terminus of 53BP1.

Figure 4.3: 53BP1: 3xFLAG Pull Down (continued)

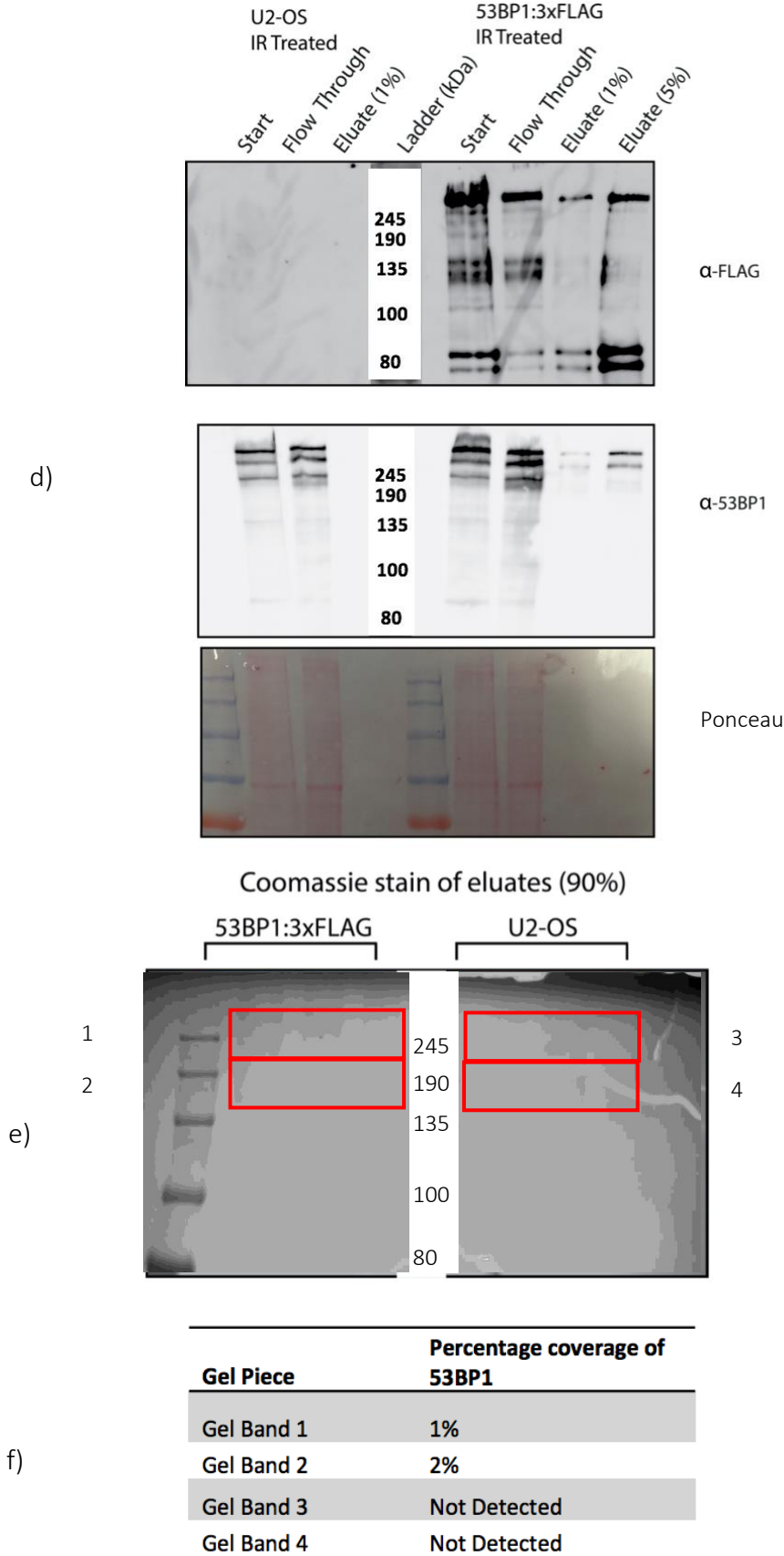


Figure 4.3 Legend: 53BP1: 3xFLAG Pull Down (continued)

- d) Western blot of FLAG tag IP. Untransfected U2-OS and stably-transfected 53BP1:3xFLAG U2-OS monoclonal cells were irradiated with 20 Gy ionizing radiation. The FLAG tag was probed to confirm transfection and to verify that the signal in the eluate was exclusive to the tagged version of 53BP1 as opposed to the endogenous 53BP1. 53BP1 was also blotted as a control for the correct protein being pulled down. Ponceau staining used as a loading control.
- e) Coomassie stained gel of the eluates. 90% of the eluate from the 53BP1:3xFLAG IP was run on a 7.5% SDS-PAGE to determine if 53BP1:3xFLAG would be detectable by MS. Gel Band 1 corresponds to proteins larger than 245 kDa and Gel Band 2 corresponds to proteins larger than 190 kDa and less than 245 kDa, as determined by the size marker.
- f) Running a data dependent method on the LTQ Orbitrap, we were only able to detect a combined 3% coverage of 53BP1 from both gel bands. We did not detect any 53BP1 from the FLAG IP used in the untransfected cell lines as a negative control.

At first, 53BP1-3xFLAG was IP'd from 2.0×10^7 cells that were irradiated with 20 Gy γ -radiation. FLAG could not be detected via western blot in the eluate. Subsequently, the number of cells were increased ten-fold and the 53BP1-3xFLAG IP was repeated using an α -FLAG antibody. Consequently, both 53BP1 and FLAG were detected in the stable cell line with western blot: 53BP1-3xFLAG was successfully enriched (**Figure 4.2d**). Bands with a molecular weight that corresponded to 53BP1 were detected when 90% of the eluates were loaded onto Coomassie stained SDS-PAGE gels. However, the bands were barely visible (**Figure 4.2e**). It was encouraging to see that there was no 53BP1 detected with mass spectrometry from the untransfected U2-OS cell line. Even though 53BP1 was detected via mass spectrometry in the stable cell line, it was only at 1-2% coverage of the peptide, which would not be sufficient for a proteomics screen.

The low peptide coverage of 53BP1 could be due to various reasons. It was not a mass spectrometry technical issue as other proteins could be detected with good signal and decent coverage as outlined in other sections. Therefore, it was hypothesised that the low peptide coverage could be due to biological reasons. The two factors that could be responsible were the promoter that was chosen, i.e. the cytomegalovirus (CMV), and the endogenous 53BP1 outcompeting the FLAG-tagged version of 53BP1. The CMV promoter has been documented as being susceptible to DNA methylation (**Grassi 2003**). Cytosine-179 from the transcription start site has been shown to be a common methylation site. Moritz et al. (2015) carried out numerous cytosine to guanine point mutations across the promoter in a plasmid containing the CMV promoter. They transfected the plasmids into Chinese hamster ovary cells. They demonstrated that C-41G and C-179G point mutations

had led to both increased protein production and extended maintenance of protein expression in stably transfected cell lines (Moritz et al. 2015).

Grassi et al. (2003) chose another route to reactivate the CMV promoter. 5'-Aza 2'-deoxycytidine (5'-Aza) has shown to be an inhibitor of DNA methylation. Its use has been shown to reactivate gene expression and potentially alter cellular differentiation. (Jüttermann et al. 1994). Trichostatin-A (TSA) has also been shown to activate gene expression by inhibiting HDACs (Mogal & Abdulkadir 2006). Grassi et al. had transfected cells with a hygromycin resistance gene that was regulated by the CMV promoter and also had GFP as a marker of expression that would be detectable by microscopy and flow cytometry. They used both of these inhibitors individually and in combination to demonstrate that they could prolong the expression of the hygromycin gene that had its expression regulated by the CMV promoter. 5'-Aza had more of an effect compared to TSA alone, but when used in conjunction, it resulted in an additive effect of increased CMV expression (Grassi 2003). The issue with these different approaches to reactivating a putatively silenced CMV promoter is that the former would require more weeks if not months for fine-tuning, and it is another point of failure in a molecular biological sense. The use of inhibitors for chromatin modifiers has an experimental caveat; the intention is to study the effects of DNA damage on HPTMs. There is not going to be any accuracy in the proteomic screen if the DOE results in manipulation of the chromatin outside of the scope of DNA damage. Also, where exactly the inhibitors would affect gene expression could not be controlled. The effect of TSA and 5'-Aza would be at a global level and not local because the inhibitor cannot be controlled from pervading throughout the nucleus.

Another hindrance to the 53BP1-3xFLAG IP system would be that 53BP1 would have to compete with endogenous 53BP1. Western blot from Figure 4.2a shows that there might be a small increase of 53BP1 expression in FLAG-tagged compared to untransfected U2-OS cell line, as the same amount of protein was loaded onto the gel. The minute increase could be a result of the total sum of FLAG-tagged 53BP1 in addition to native 53BP1 in the transfected cell line. If the difference were to be subtracted, it would show that 53BP1-3xFLAG's expression is a small fraction compared to the native 53BP1 regarding total 53BP1 expression. An option one could consider to potentially increase the yield of 53BP1-3xFLAG during IP is to completely replace the endogenous version of 53BP1 with the FLAG-tagged version. In human cells, 53BP1 is located on chromosome 15 (The UniProt Consortium 2017). Canonically, somatic cells possess only two copies of chromosome 15. However, some osteosarcoma cell lines have been shown to possess multiple copies of chromosome 15; aneuploidy and polyploidy are not uncommon in cancerous cells (Ozaki et al. 2003). Another issue with using cancerous cell lines is that their expression levels of a given gene have deviated from the "wild type" and can be stochastic. Therefore, it becomes imperative that if one wishes to use 53BP1-3xFLAG to pull down DNA damaged chromatin one should mitigate against the intracellular competition of what one is trying to achieve by completely replacing the endogenous version of 53BP1 with 53BP1-3xFLAG. However, the best use of carrying out gene editing would be with the use of CRISPR-Cas9 editing system as compared to TALEN based systems, CRISPR-Cas9 requires less effort to tailor the system to new desired DNA sequences and that cleavage takes place at a more specific site rather (between the 17th and 18th of the target sequence) have the cleavage takes place in a non-specific manner,

where it ranges from 12-24 bp as seen in a TALEN system (Ran et al. 2013). An improvement to this approach in the future, would be to implement a modified CMV promoter so that it does not become methylated in the future as demonstrated by Moritz et al. (Moritz et al. 2015), which would take several weeks with an optimistic timescale. Also, the implementation of knocking-out the endogenous 53BP1 and replacing with the tagged version would also take several weeks if not months. So, even though the 53BP1-3xFLAG showed glimpses of working, it would require a significant overhaul of the cell line to get working with possible efficiency. From here on, it was decided to focus solely on γ H2AX IP as it was showing to be more promising.

4.2.2. Lymphoblastoid for DNA Damage Proteomic Screen

Pulling down γ H2AX directly with a specific antibody was the other option for enriching DNA damaged-related chromatin. Three antibodies for γ H2AX were selected: Millipore clone JBW301 05-636, Millipore clone JBW301 16-193, and Abcam (ab2893). The three antibodies were used to immunoprecipitate γ H2AX from the same lymphoblastoid cell line (LCL) sample that was irradiated with 20 Gy of IR. The Abcam antibody had enriched the most γ H2AX from the LCL sample (Figure 4.4a). A technical replicate was carried out with a new biological sample to ensure the result was accurate. The JBW301 16-193 antibody has a biotin conjugate, therefore γ H2AX enrichment with streptavidin was also attempted to determine if it was better able to enrich compared to Dyna IgG beads. Again, the Abcam antibody was the most efficient at pulling down γ H2AX (Figure 4.4b).

Figure 4.4: Antibody Selection for γ H2AX IP

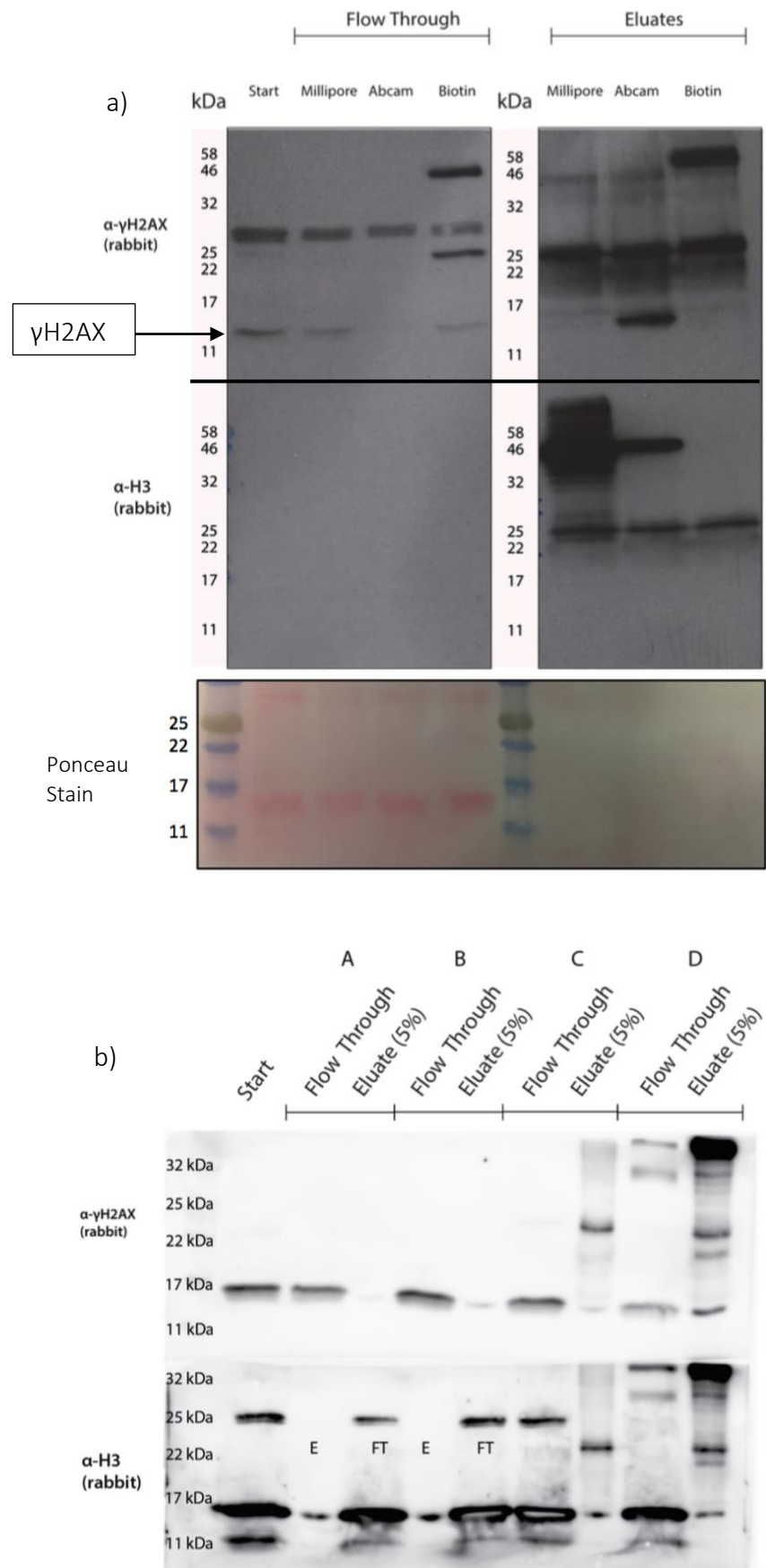


Figure 4.4 Legend: Antibody Selection for γ H2AX IP

- a) Western blot of γ H2AX IP after the use of various antibodies to immunoprecipitate γ H2AX. Lymphoblastoid cell line (LCL) were irradiated with 20 Gy ionizing radiation. 200 μ g of digested nucleosomes from the same biological sample were used for each IP. Three different antibodies were used for the IP to assess which antibody gives the best yield of γ H2AX: Millipore (JBW301), Abcam (ab2893 Chip grade), and Biotin (JBW301-biotin conjugate). α - γ H2AX was blotted to determine yield of γ H2AX. Ponceau staining and α -H3 loading controls failed but Ponceau indicated similar loading of protein.
- b) Western blot of biological replicate γ H2AX IP with various antibodies to immunoprecipitate γ H2AX. LCLs were irradiated with 20 Gy ionizing radiation. 200 μ g of digested nucleosomes from the same biological sample were used for each IP. Three different antibodies were used for the assay to assess which antibody gives the best yield of γ H2AX: Millipore (JBW301), Millipore (JBW301-biotin conjugate) and Abcam (ab2893 Chip grade). α - γ H2AX was blotted to determine yield of γ H2AX. α -H3 was blotted as a loading control.

A: Millipore, clone JBW301 05-636, with strep beads

B: Millipore, clone JBW301 16-193, biotin with dyna (IgG) beads

C: Millipore, clone JBW301 16-193, biotin with with strep beads

D: Abcam (ab2893) with dyna (IgG) beads

E: Elution

FT: Flow Through

It was decided to use LCLs as opposed to U2-OS cells for the γ H2AX-IP-MS. U2-OS cells are more suited for engineering stable cell lines. However, with the γ H2AX-IP-MS, there would be no need to introduce any tags since the different tail of H2AX acts as an intrinsic tag; mainly when an antibody that only recognises the phosphorylated form of H2AX is used. LCLs have more practical advantages compared to U2-OS cells: they are non-adherent, so they grow in suspension. Thus more cells can be grown within a given space compared to U2-OS that are adherent. Their lack of adsorptive nature makes them easier to handle when inducing DNA damage via IR. Standard flasks that are used for growing cells cannot fit inside the IR-source, so the cells require being transferred to a 50 mL conical centrifuge tube. This is not such an issue for LCLs as they can be decanted in sterile conditions. However, U2-OS cells require treatment with trypsin to loosen them from the surface of cell culture flask. The concern is that unintended cell responses are being introduced when they are not in their natural state, e.g. hypoxia could introduce a DNA damage response (Olcina & Hammond 2013). Another issue would be how the cells cope with recovery: U2-OS cells would have to simultaneously be dealing with adhering to the surface while also responding to the DNA damage. LCLs would only be concerned with repairing their IR-induced DNA breaks. Another molecular advantage of using LCLs compared to U2-OS is that as long as LCLs have not undergone many rounds of passages then the IR induced DNA damage response should have a better representation of what takes place in regards to DDR-induced HPTM in mammalian physiology as compared to U2-OS cells (Sie et al. 2009). LCLs have been used in other labs in regards to the transcriptional response of DNA damage, where 3 Gy to 10 Gy of IR were used to challenge the cells (Jen & Cheung 2003). They have also been used in proteomic studies

to determine the responsiveness of DNA damage response proteins after treatment with bleomycin (Dirksen et al. 2006).

4.2.3. Native or Crosslinking for γ H2AX IP

Another question that arose is whether to use a crosslinking agent for the immunoprecipitation (IP) of proteins. While crosslinking-chromatin immunoprecipitation (xChIP) has its advantages over native ChIP (nChIP), those advantages would not apply to immunoprecipitating nucleosomes (Turner 2001). xChIP is useful when one wants to assess the interactive partners of a protein complex which would involve proteins that possess weak affinity to DNA or do not interact with DNA directly. Fixing the proteins in place would lead to a snapshot of the proteins present at a specific site at a given moment of time, which is useful for determining how a particular reagent or gene deletion affects the molecular machinery involved in a given biological process. Another advantage to this would be that the chromatin remained fixed during the preparation and immunoprecipitation stages. However, the benefit of keeping chromatin fixed in place does not apply to γ H2AX-IP-MS. This is because the interaction of DNA and histones is so strong that they form nucleosomes in vitro. This interaction was characterised in the 1970s (Axel et al. 1974), (Oudet et al. 1975), (Laskey et al. 1978), and (Nelson et al. 1979). Also, pulling down extra nascent protein could increase the background noise and reduce the efficacy of detecting HPTM changes. Another issue with crosslinking is that it requires an incubation step which can affect time sensitive data collection; time points were collected mere minutes after the induction of DNA damage, and most crosslinking protocols require an incubation step that is even longer than the shortest recovery in the DOE (Pchelintsev et al. 2016). However, it would be less of an issue when assessing HPTMs at late repairing breaks. The other disadvantage with cross-linking is that it

requires a sonication step to shear the fixed chromatin. The issue with this is that sonication can be imprecise and lead to lower yields of both DNA and protein. It has been shown with the use of electron microscopy that sonication can alter the structure of the protein. Their structure exhibits properties similar to that of amyloids (Stathopoulos et al. 2008). Alternative methods have been developed that make light use of sonication (Pchelintsev et al. 2016). However, it would be most efficient to abstain from it in this case. Tseng et al. (2014) developed an H2AX purification protocol that enables the mapping of where particular histone variants were located in the genome of embryonic stem cells; in this case, they had avoided the use of crosslinking and used native ChIP-seq (Tseng et al. 2014). They were then able to demonstrate that H2AX deposition in embryonic stem cells works as a signal for identifying the developmental potential of pluripotent stem cells and H2AX downregulates the expression of extraembryonic genes in pluripotent stem cells (Wu et al. 2014). David et al. (2017) cross-compared the same sample of chicken pectoral tissue that was either crosslinked and not crosslinked. They found that after sequencing, the sample that had been processed with nChIP had given a much higher hit rate of identifying H3K27me3 regions compared to the tissue sample that had undergone xChIP: 11,008 regions vs 611 regions for nChIP-seq and xChIP-seq, respectively. They were then able to confirm the discrepancies between native and crosslinking with n/x-ChIP-qPCR. (David et al. 2017). Therefore, to ensure the highest possible yield of pulling down chromatin, native ChIP would be the best approach for γ H2AX-IP-MS.

4.2.4. Antibody Selection for γ H2AX IP

There is a vast library of γ H2AX antibodies available from over a dozen manufacturers.

Therefore, with the plethora of antibodies raised for γ H2AX, it was paramount to not only choose the right antibody for the application, but the antibody's specificity for the chosen antigen would need to be tested and its use in immunoprecipitating γ H2AX was reproducible. Linscott's Directory is a search engine for antibodies that trawls the catalogues of various biotech manufacturers. If one were to search for γ H2AX, it would return a result of 175 antibodies from twelve suppliers. The result can be filtered down to a desired specification: monoclonal or polyclonal; host species; reactive species; conjugated or not; and if available the known applications where the antibody is useful in. Linscott's Directory does not even include results from Abcam and Sigma Aldrich, which would push the results potentially north of 200 results. Manually characterising 200 antibodies is neither feasible or practical. Therefore, a select few antibodies were shortlisted: some were available in the lab stock that were used for other applications and others were procured for testing for the assay.

The three antibodies tested here were: Millipore clone JBW301 05-636, Millipore clone JBW301 16-193, and Abcam (ab2893). To validate the shortlisted antibodies, the Rimm Lab Algorithm for antibody validation was adapted and modified. The algorithm works by employing sequential selective control tests to ensure that a given antibody is reliable, specific, and suitable for a given assay (**Bordeaux et al. 2010**). Rimm's algorithm was initially designed for immunohistochemistry and quantitative immunofluorescence. In this case, the methodology was tailored for proteomics (**Figure 4.5**). As mentioned, the vast numbers of antibodies available for γ H2AX would have unintended consequences of

potentially choosing the wrong antibody for immunoprecipitating. The procedure requires identification of several suitable γ H2AX antibodies candidates from several manufacturers, which had been already carried out. LCLs were irradiated to induce damage and carry out a western blot to determine that the antibodies can recognise the S139phos epitope. The bands were checked to ensure that they correspond to the molecular weight of the protein of interest; in this case, it would be around 15 kDa for the unubiquitinated version. If the size of the bands does not correspond to the expected size, the original algorithm suggests the next step would be to follow through with IP, but γ H2AX is such a powerful signal that if there are no bands or multiple bands for a protein, then the antibody would be discarded. If the molecular weight is what was expected, then an IP for γ H2AX with a single time point would be the next step. A positive and negative control cell line were included as well to give further credence to the antibody. Once the negative and positive results illustrated that the antibody was functional for the task, the next step was to titrate the antibody for optimal salt concentrations and determining what is the least amount of antibody that can be used that results in depletion of γ H2AX from the lysate. The cells were irradiated with 20 Gy of IR. Even though such a large dosage of radiation is not physiologically relevant, as 2-8 Gy is lethal, it would ensure that the cells give a tremendous DNA damage response, resulting in more histones undergoing post-translational modifications and in turn potentially producing large yields of nucleosomes from damaged chromatin during the immunoprecipitating. One needs to take into consideration, that even though H2AX accounts for approximately 5% of all H2A variants (Chapter 3), not all of the H2AX will be phosphorylated: the number of nucleosomes that are pulled down could be minuscule when compared to all the chromatin. The Abcam antibody resulted in the most enrichment of γ H2AX: it was almost

entirely depleted from the flow-through. Comparitively, both the biotinylated and unbiotinylated Millipore antibodies had negligible differences in their efficiency of pulling down γ H2AX. However, compared to Abcam there was a significant reduction in γ H2AX in the eluate (**Figure 4.4a**). A technical replicate was repeated to ensure the result was not an anomaly, and an affinity purification was attempted with the biotin-conjugated Millipore antibody by using streptavidin beads instead of IgG beads. Just in case there was a chance the affinity purification resulted in a higher yield of enriched γ H2AX. However, a repeat result was observed: the Abcam antibody was much more suited to this application compared to the other antibodies, regardless of using IgG or strep beads. (**Figure 4.4b**). Now that the candidate antibody had been filtered to one, the method of getting the γ H2AX-IP-MS to work was only beginning, as there were further quality control checks that needed to be satisfied before being fully deployed.

Figure 4.5: DOE Algorithm for antibody validation for the use of γ H2AX IP-MS

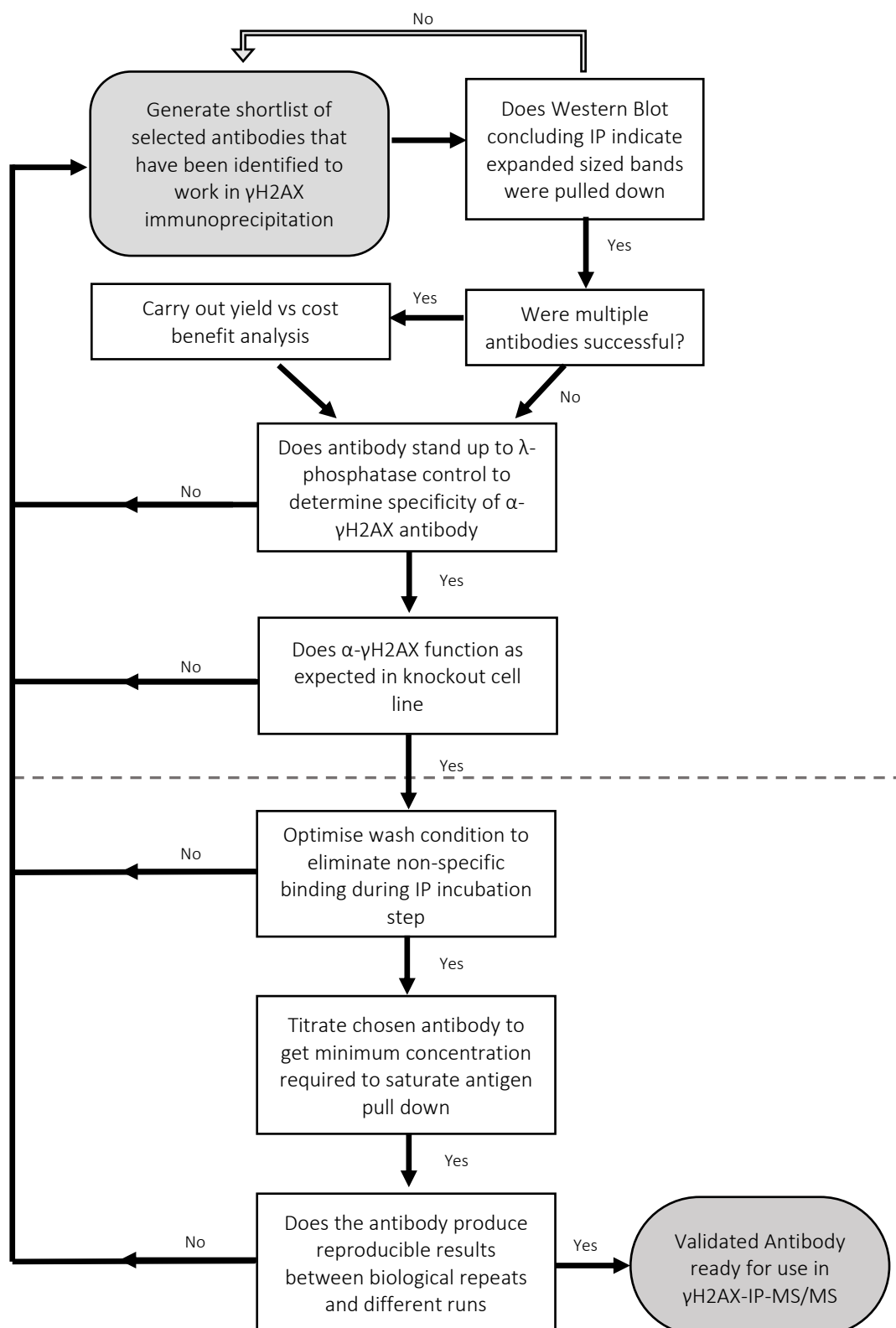


Figure 4.5 Legend DOE Algorithm for antibody validation for the use of γ H2AX IP-MS

(above the dotted line) Step 1 of antibody validation involves selecting the best candidate with a sequential gauntlet which involves a direct multivariate IP, followed by using several negative controls: (λ -phosphatase and MEF H2AX^{-/-})

(below the dotted line) Further validation; optimisation and reproducibility of γ H2AX antibody on isolated nuclei for protein of interest localization. Finalised with reproducibility between biological repeats and different antibody lots.

4.2.5 γ H2AX IP optimisation: λ -Phosphatase Treatment

To validate the antibody further, some negative controls were employed that challenged the fidelity of the chosen γ H2AX antibody's ability to bind to its epitope. It was hypothesised that if the antibody were only binding to phosphorylated S139 of H2AX, the antibody would not enrich for any proteins after treatment with a phosphatase. Even in the presence of large amounts of damage. Lambda protein phosphatase (λ -phosphatase) was originally discovered in bacteriophage λ by Cohen & Cohen (1989) (Cohen & Cohen 1989) where at the time it was putatively characterised to be a phosphatase that was dependent on Mn^{2+} ions. Further studies of λ -phosphatase by Zhuo et al. (1993) (Zhuo et al. 1993) confirmed its ability to remove the phosphate modification from serine, threonine and tyrosine.

The rate of dephosphorylation varied with different proteins. They also characterise the biochemical activation and inhibition by metal and non-metal ions, respectively. To minimise any discrepancy between any samples, the same biological material was used throughout: a harvest of LCL was divided in two with a ratio of 1:2. The larger portion was treated with 20 Gy of ionising radiation, and the smaller portion was mock irradiated. After recovery, the IR treated cells were split in two again but this time equally. One subset of the IR treated cells were incubated with λ -phosphatase with the intention of removing the phosphorylation from γ H2AX. Thus ending up with three samples of LCL: mock treated, IR-treated and λ -treated, and only IR-treated sample. The antibody of choice was further certified as suitable for the task when another γ H2AX-IP was carried out with the three samples (Figure 4.6a). The IR treated sample had enriched γ H2AX as expected. However, encouraging results came from no γ H2AX being enriched in the λ -

phosphatase treated sample, despite those same cells being treated with the same gamma-source simultaneously as the IR-only treated sample. The levels of γ H2AX signal in the λ -phosphatase treated cells were at similar levels to that of mock-treated cells. The lack of detectable signals was further corroborated by the fact that barely any histones were visible when the eluates were run on an SDS-PAGE (**Figure 4.6a**). From this exercise, two results were obtained: the choice of antibody was highly specific and thus more validated, and that λ -phosphatase treatment is both a thorough and contextual control step to γ H2AX and should be employed for future γ H2AX-MS-IP assays.

Figure 4.6: γ H2AX IP optimisation: λ -Phosphatase Treatment

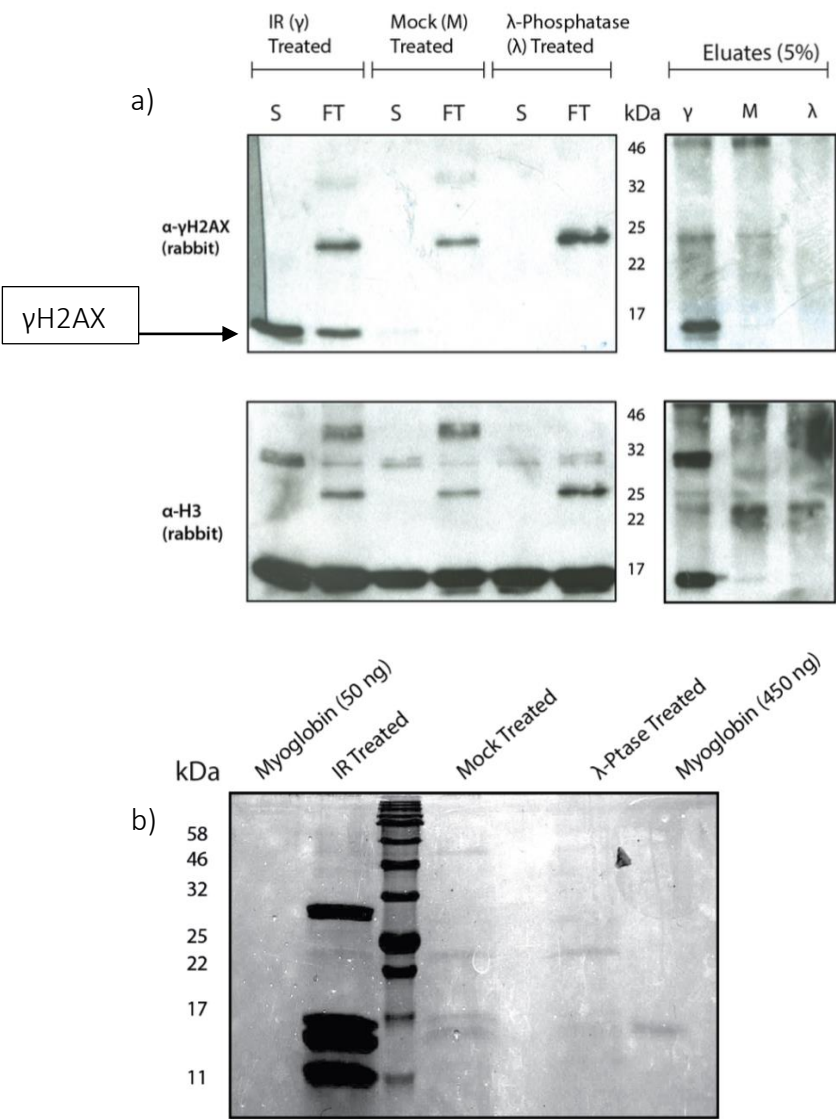


Figure 4.6 Legend: γ H2AX IP optimisation: λ -Phosphatase Treatment

- a) Western blot of γ H2AX IP with three different conditions for measuring specificity of the γ H2AX IP using the antibody from Abcam. LCLs were irradiated with 20 Gy ionizing radiation and a mock sample that was not treated to ionizing radiation were digested with micrococcal nuclease. 1 mg of digest histone sample was used for each IP sample. Prior to incubating the antibody with the sample, 1 mg of 20 Gy LCL irradiated digested histones was treated with λ -phosphatase to remove the phosphorylation modification from S139. λ -phosphatase used: New England BioLabs P0753.
- b) Coomassie stained 15% SDS-PAGE gel of the eluates. 30% of the eluate from the γ H2AX IP was run to determine yield of nucleosomes enrichment.

S; Start

FT; Flow-Through

4.3.8 MEFS as control for antibody specificity

To further verify that the antibody was only pulling down H2AX phosphorylated at S139 and not another protein that was also phosphorylated with a coincidentally similar migration, the Abcam antibody was used for another immunoprecipitation in a mouse embryonic fibroblast cell line that did not express H2AX (MEF H2AX^{-/-}). The argument could be made that the λ -phosphatase treatment just shows that the antibody binds to a phosphorylation site, but it does not specify which one. It is still possible that the antibody could be pulling down proteins that possess phosphorylation sites that act as undesired epitope sites and happen to be similar in size to a histone in kDa. As the Rimm Lab Algorithm was adapted for antibody validation, the next step would be to try the same experiment in a negative control cell line. In this case, the negative control would be γ H2AX knockouts. Mouse and human H2AX possess very strong homology (**Figure 4.7b**). Also, mouse genome having undergone extensive genomic rearrangement (**Graphodatsky et al. 2011**). Despite this, and the cells not having human origins, it would not be problematic; the DNA sequences of the genes of humans and mice remain very homologous. In the case of H2AX, which is the protein of interest, mice isoforms of H2AX possesses a difference of only four amino acids when contrasted against human H2AX sequence. In spite of the strong homology, a MEF cell line that did express H2AX was used as a positive control for the negative control, to illustrate that if the antibody did not pulldown any histones from the MEF H2AX^{-/-} cell lines, it is not because it derived from a different species. LCLs, MEF H2AX^{+/+} and MEF H2AX^{-/-} were challenged with 20 Gy IR. In addition, LCLs were also mock-treated. The Abcam antibody was able to enrich γ H2AX from both the irradiated LCLs and the H2AX^{+/+} MEFs. However, no γ H2AX was enriched

with either the mock-treated LCLs or H2AX^{-/-} MEFs, suggesting that the antibody was highly specific for γ H2AX and suitable for the IP of damaged chromatin.

Figure 4.7: MEFS as control for antibody specificity

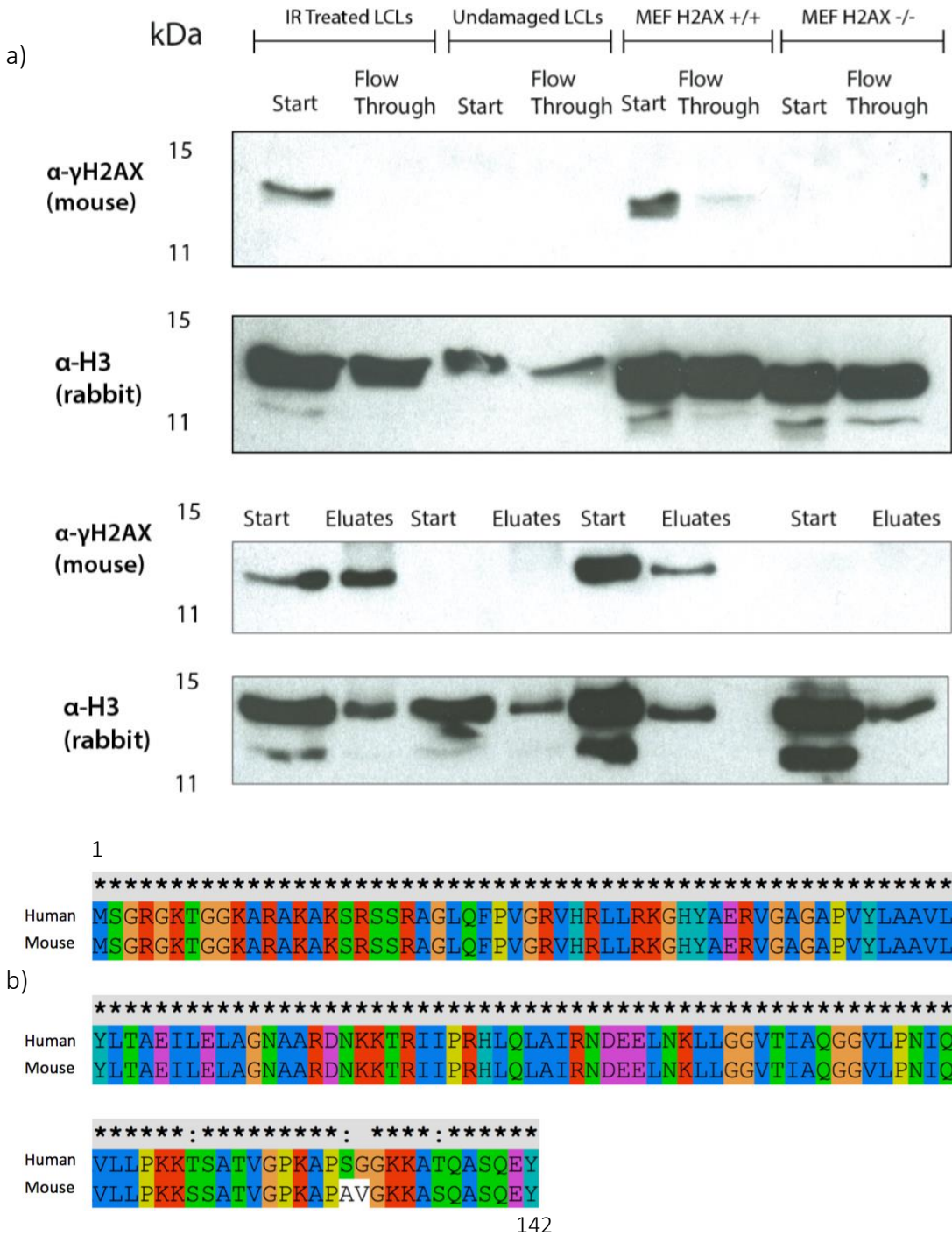


Figure 4.7 Legend: MEFS as control for antibody specificity

- a) Western blot of γ H2AX IP with three different cell lines for measuring specificity of the γ H2AX antibody from Abcam. LCLs, MEFs and MEFS H2AX^{-/-} were irradiated with 20 Gy ionizing radiation. 1 mg of Mnase-digested histone sample was used for each IP sample. MEF H2AX^{-/-} were used as a negative control for the γ H2AX antibody. Wild type MEFs were used as positive control for the γ H2AX antibody binding to the mouse version of H2AX.
- b) Sequence homology of human H2AX and mouse H2AX. The two mammalian H2AX isoforms have a four amino acid residue difference to each other.

Not only are the function of HPTMs conserved amongst mammals but in which cell they take place is also conserved (Woo & W. H. Li 2012). Therefore, in this scenario, it would be optimal to include a wildtype MEF that expressed H2AX in acting as a control for the negative control, i.e. the MEF H2AX^{-/-}. At this stage, it would not cause the antibody to fail to recognise H2AX just because the cells are from a different mammalian species. Subsequently, LCLs, MEF H2AX^{+/+} and MEF H2AX^{-/-} were irradiated with 20 Gy of ionising radiation. There was also a batch of LCLs that were mocked-treated. When the γ H2AX-IP was carried out, enrichment of γ H2AX was detected in the IR treated LCLs and the MEF H2AX^{+/+} cells (**Figure 4.4b**). The fact that the antibody was able to recognise γ H2AX across species, and not recognise a false positive when H2AX was either not phosphorylated or not present, i.e. mock treated or λ -phosphatase treated and MEF H2AX^{-/-}, respectively, made the antibody suited for the task of enriching chromatin that is proximal to DNA damage. However, in the negative samples, it happened to be that histone H3 was being pulled down, which was visible in all of the eluates (**Figure 4.7a**). The abundance of H3 signal suggested that there was non-specific binding, which could affect proteomic results, so that would also require further optimisation. However, at this point, the antibody could not be validated any further in a practical manner and thus completed the adapted algorithm for validating the antibody and for its practical purpose.

4.2.6 IP of γ H2AX with stringent conditions

Even though γ H2AX was not detected in the eluates of the negative controls, histone H3 was still detected in the eluates. It was reasoned that increasing the washing and including detergents could address this issue. The number of washes was increased from three to five; two different detergents were added, NP40 and Tween-20, at concentrations of 0.1%; and finally, the wash buffer volume was increased from 400 μ L to 800 μ L. The issue of non-specific histone H3 being pulled down in the negative control was resolved (**Figure 4.8a**), but the addition of detergent did not make a difference to the non-specific histone H3 that was enriched with γ H2AX.

Detecting histone-H3 in the eluates of the negative controls but not γ H2AX itself heavily implied that there was a non-specific binding problem (**Figure 4.7a**). The concern with the wash step is that there is a balance between retaining as much of the protein as possible and excluding as much background noise as possible. For example, if the salt concentration is too high, then it might lead to premature elution of the protein. However, if the salt concentration is too low, unwanted non-specific protein will not be ejected which could hinder both the signal-noise ratio and interpreting results. Since there was non-specific binding, an increase in the harshness of the wash step was required; to avoid manipulating the different salt concentrations as that would require more elaborate optimisations (Chapter 2) for the whole IP step, compared to tinkering with the frequency of washes. The first alteration was to increase the number of washes from three to five. The volume of the wash buffer was increased from 400 μ L to 800 μ L. The rationale for both of these alterations was simple: the larger the volume of wash

buffer the beads are exposed to the more likelihood of washing away weak non-specific bound proteins. The other alteration was to use detergent (0.1% NP40 and 0.1% Tween-20) which should decrease non-specific binding. The γ H2AX-IP was repeated with the prescribed wash conditions, and when comparing detergent based washing vs non-detergent based washes, no difference in non-specific binding was observed (**Figure 4.8a**). However, it was decided the detergent in the wash step was kept in the protocol as it has the serendipitous effect of making the solution less viscous, which resulted in decreased surface tension and so there were less magnetic beads lost with each subsequent wash cycle. There was, however, an immediate and significant difference in the reduction of histone H3 in negative control eluate. This result concurred with the hypothesis that both free histones and octomers would bind non-specifically as they have a strong affinity for DNA, especially double-stranded blunt-ended DNA (**Almouzni et al. 1990**). Therefore, it was imperative to eliminate any non-specific chromatin as it could affect the results of the proteomic screen. With the wash step optimised, the samples were ran on a SDS-PAGE gel, which was stained with Coomassie. There were no visible signs of histones in the negative control (**Figure 4.8b**). This was a very positive result as histones are available in abundance in the genome, and their cluster which spans from 9 kDa to 17 kDa can be detected quite easily with Coomassie staining of an SDS-PAGE gel. The fact the lanes on the gel that correspond to the negative control were almost clear indicates that the modification in the wash step resulted in high efficacy in eliminating non-specific bound protein.

The detergent was still retained in the wash step as it reduced the surface tension which had the positive consequence of losing fewer beads with each wash step. The Coomassie

stain showed an abundant enrichment of histones in the eluate of LCLs and MEF H2AX^{+/+} compared to MEF H2AX^{-/-}, where there was virtually none (**Figure 4.8b**). The input and eluate of the LCL sample was prepared for mass spectrometry. A ten-fold increase in the enrichment of γ H2AX as a proportion of all H2AX was observed. Another technical replicate was carried out to ensure the wash steps were efficient in removing non-specific nucleosomes from the eluate (**Figure 4.8c**).

Figure 4.8 IP of γ H2AX with stringent conditions

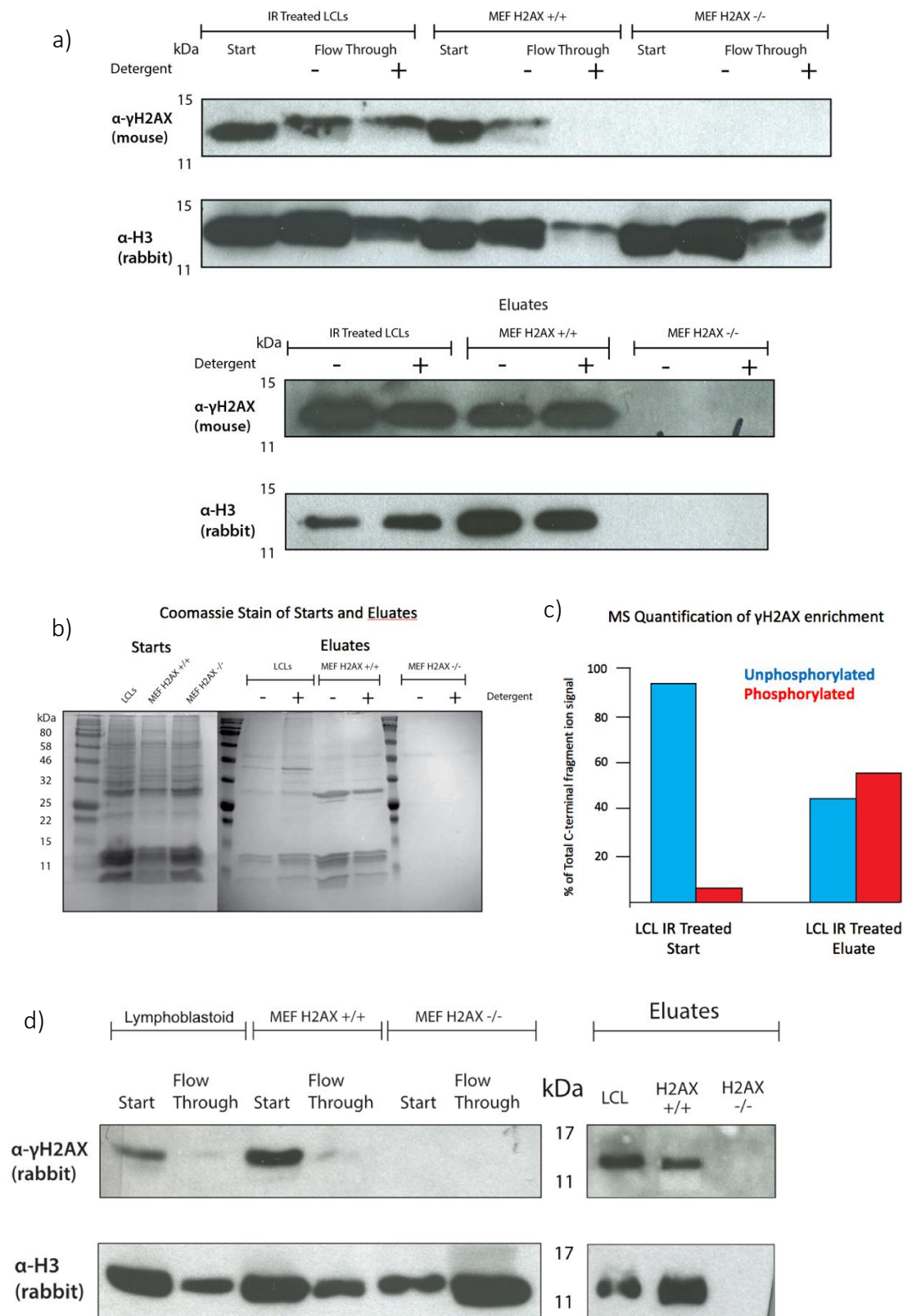


Figure 4.8 Legend: IP of γ H2AX with stringent conditions

- a) Western blot of γ H2AX IP with three different cell lines for measuring specificity of the γ H2AX antibody from Abcam. LCLs, MEFs and MEFS H2AX^{-/-} were irradiated with 20 Gy ionizing radiation. 1 mg of Mnase-digested histone sample was used for each IP sample. MEF H2AX^{-/-} were used as a negative control for the γ H2AX antibody. Wild type MEFs were used as positive control for the γ H2AX antibody binding to the mouse version of H2AX. Each sample was incubated with the γ H2AX antibody and with and without the presence detergent (0.1% NP40 and 0.1% Tween-20). Number of washes were increased from three to five. Volume of wash buffer was increased from 400 μ L to 800 μ L
- b) Coomassie stained SDS-PAGE gel of the eluates. 30% of the eluate from the γ H2AX IP was run to determine yield of nucleosomes enrichment. 15% SDS-PAGE gel
- c) MS quantification of the γ H2AX enrichment. Phosphorylated peptide increased from 5% in the start to approximately 60% in the eluate. Singly charged fragment ions from ATQASQEY and ATQAS[p]QEY precursor ions were used for quantification.
- d) western blot of γ H2AX IP with three different cell lines for measuring specificity of the γ H2AX antibody from Abcam. LCLs, MEFs and MEFS H2AX^{-/-} were irradiated with 20 Gy ionizing radiation. IP without incubation with detergent. 200 μ g of digested nucleosomes were used for each IP sample.

4.2.7 Antibody Titration

The last step in the antibody optimisation was a titration to determine the minimal amount of antibody required to deplete sufficient quantities of γ H2AX from the extracted nuclei. The crass option would be to use as much antibody as possible, but again, to ensure that the project remained as economically feasible as possible various concentrations were tested: 0.5 μ g/mg of protein to 10 μ g/mg which was the manufacturers recommended concentration. The optimal concentration was found to be around 2.5-5 μ g of antibody per mg of protein and any concentration above that would result in diminishing returns (**Figure 4.9**). Histone H3 was not depleted with the highest concentration of antibody due to the nature of H2AX only occupying a small proportion of nucleosomes compared to the rest of the genome.

Figure 4.9: Antibody Titration

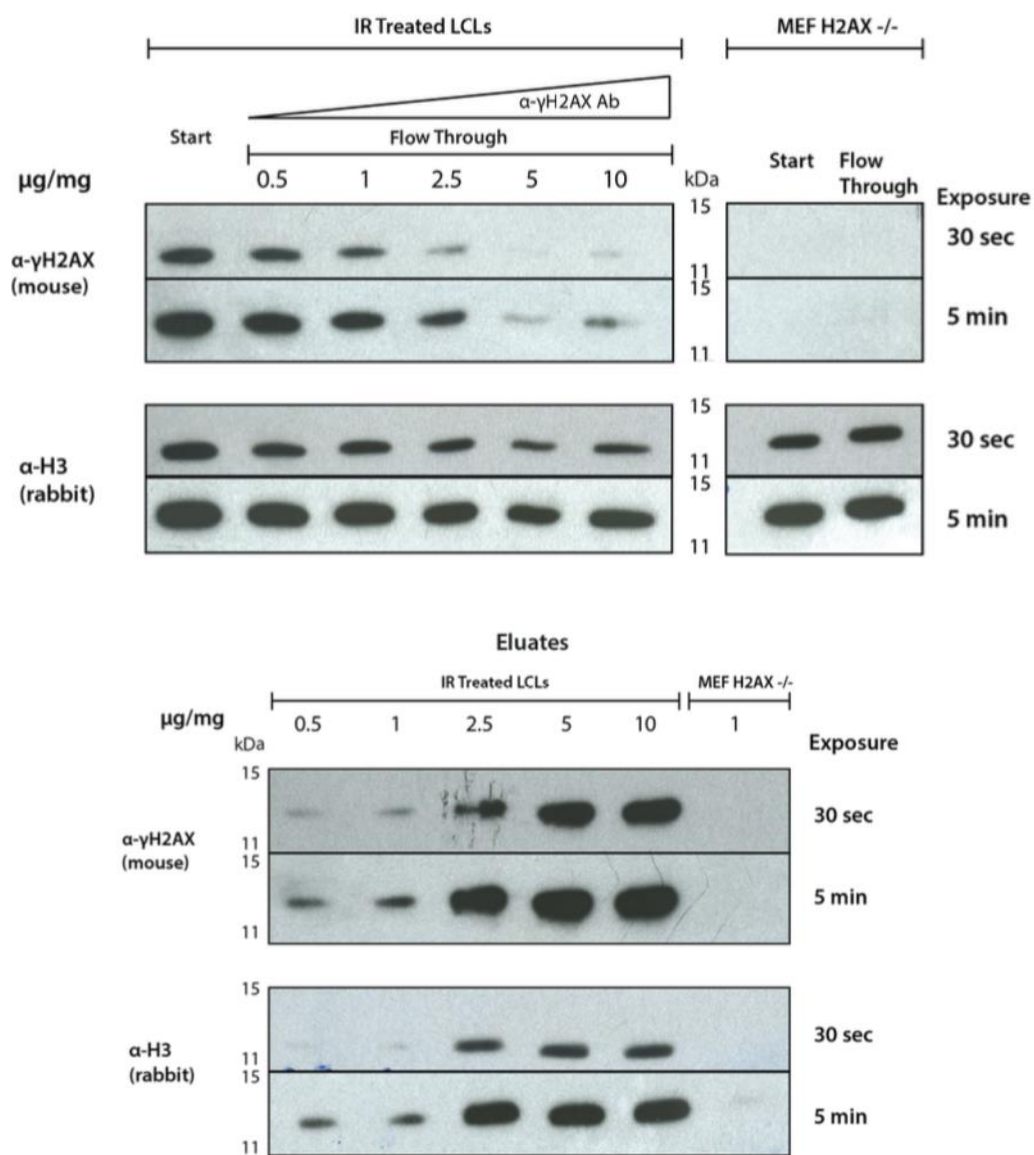


Figure 4.9 Legend: Antibody Titration

Western blot of γ H2AX antibody titration with two different cell lines to determine optimal concentration of γ H2AX antibody for IP. LCLs and MEFS H2AX^{-/-} were irradiated with 20 Gy ionizing radiation. 200 μ g of histone sample was used for each IP sample. MEF H2AX^{-/-} were used as a negative control for the γ H2AX antibody. Different concentrations of antibody for IP tested were 0.5 μ g/mg, 1 μ g/mg, 2.5 μ g/mg, 5 μ g/mg, and 10 μ g/mg. 1 μ g/mg antibody concentration was used for MEF H2AX^{-/-}. μ g/mg = antibody/lysate protein

4.2.8 Mnase optimisation

To determine the efficiency with which Micrococcal Nuclease (Mnase) was digesting the chromatin, assays were conducted under different conditions to test DNA digestion. The protease inhibitors that were used during the nuclei isolation step were inhibiting the enzymatic activity of the Mnase in both the agarose gel and Coomassie stained SDS-PAGE gel (**Figure 4.10a and 4.10b**). Subsequently, a titration of the protease inhibitors was carried out with four conditions with one no Mnase condition acting as a negative control. It was also observed that the inhibitors that were used to keep proteases and intrinsic chromatin modifiers from functioning after the nucleosomes were extracted were also inhibiting the Mnase digestion of the chromatin (**Figure 4.10a and 4.10b**). Titrations were carried out that assayed the maximum concentration of inhibitor that the Mnase will tolerate before its activity is quenched. It was found that the optimal amount of inhibitor was that much lower of what is used during the nucleosome preparation (**Figure 4.10d**). As a result, the protocol was altered so that the concentration of inhibitors were decreased to below the maximum tolerance threshold that Mnase during the digestion step. Upon carrying out a titration of inhibitors and time-course to determine optimal incubation, it was determined that 50-60 minutes of treatment would result in a great yield of nucleosome chains spanning 1-2 nucleosomes in size in the S1 and approximately 5-8 nucleosomes in the S2. When there is no Mnase in the sample, there is a visible band observable at the top of the gel, highlighting that the DNA has not been digested at all. There is a correlation of longer incubation of Mnase with less smears being present and more of the DNA being found towards the bottom of the gel. The smears on the agarose

gel represent varying lengths of DNA corresponding to the bands with which they align in the DNA size ladder.

Figure 4.10 Mnase Optimisation

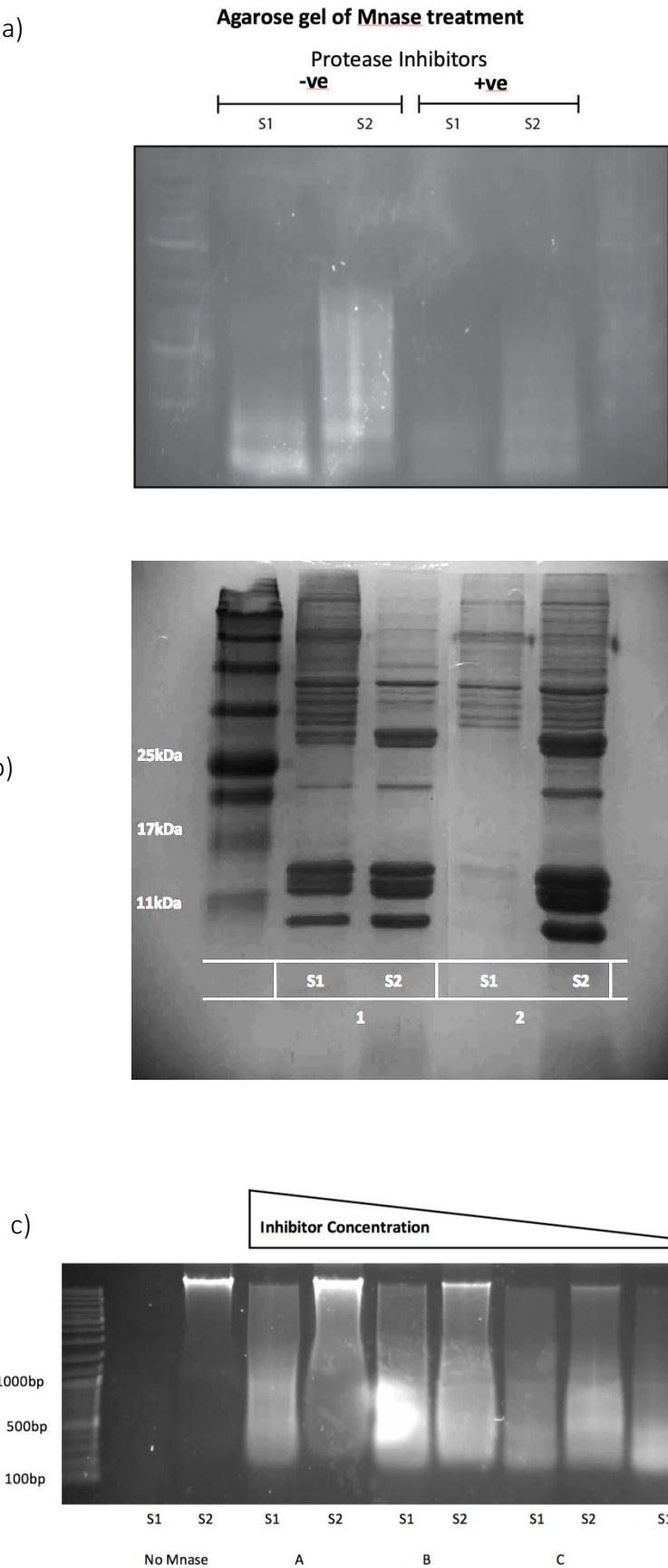


Figure 4.10 Legend: Mnase Optimisation

- a) Protease Inhibitors inhibit activity of Micrococcal nuclease. Agarose gel of Mnase treatment of LCLs. 1.5 µg of DNA was loaded onto a 1% agarose gel and ran at 50 V for 45 minutes. The same biological replicate of isolated nuclei aliquoted and were treated with two different conditions:
1. NEB Buffer for Mnase: 50mM Tris and 5mM CaCl₂.
 2. NEB + sucrose + inhibitors.
- S1: First supernatant after spin down post-Mnase treatment
S2: Second supernatant after spin down post-Mnase treatment
- b) Protease Inhibitors prevent activity of Micrococcal nuclease. Coomassie stain gel Mnase treatment of LCLs. 1.5 mg of histone was loaded onto a 15% SDS-. The same biological replicate of isolated nuclei were treated with two conditions: 1. NEB Buffer for Mnase: 50mM Tris and 5mM CaCl₂. 2. NEB + 250mM Sucrose. 3. NEB + sucrose + inhibitors
- c) Titration of inhibitors with Micrococcal Nuclease. Agarose gel of Mnase treatment of LCLs. 1.5 µg of DNA was loaded onto a 1% agarose gel and ran at 50 V for 45 minutes. The same biological replicate of isolated nuclei were treated with four conditions (all had NEB Buffer for Mnase: 50mM Tris and 5mM CaCl₂):
- i) 10 mM Na-butyrate, 0.5 mM AEBSF, 2 µM leupeptin, 1 µM Pepstatin A, 50 mM NaF, and 5 mM NEM
 - ii) 10 mM Na-butyrate, 0.5 mM AEBSF, 2 µM leupeptin, 1 µM Pepstatin A, 5 mM NaF, and 1 mM NEM
 - iii) 5 mM Na-butyrate, 0.2 mM AEBSF, 2 µM leupeptin, 1 µM Pepstatin A, 5 mM NaF, and 0.5 mM NEM
 - iv) 5 mM Na-butyrate, 0.2 mM AEBSF, 10 µM leupeptin, 1 µM Pepstatin A, 1 mM NaF, and 20 µM NEM

Figure 4.10: Mnase Optimisation (continued)

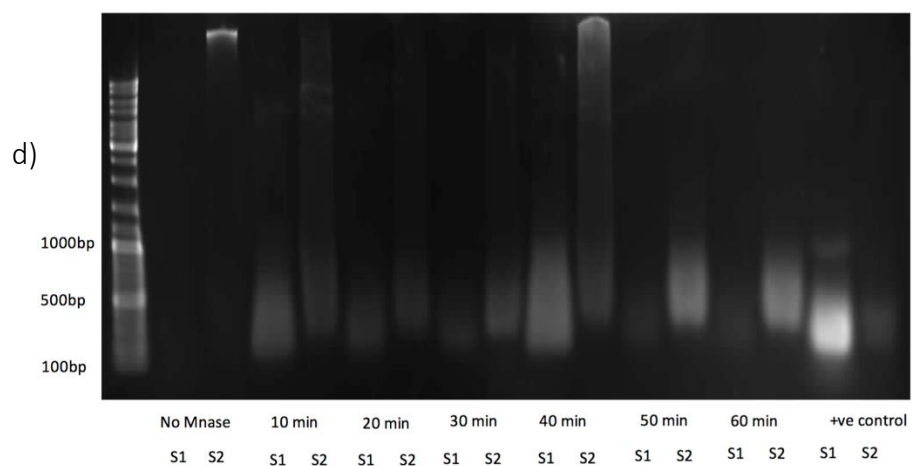


Figure 4.10 Legend: Mnase Optimisation (continued)

- d) Time course of Micrococcal Nuclease: 150 Worthing Units of Micrococcal nuclease was used per 150 μg DNA. The same nuclear isolation sample was distributed equally into seven parts. Then equal amounts of Micrococcal nuclease were added to each part and the reaction was quenched every ten minutes for up to 60 minutes. This assay was intended to see how long is required for optimal digestion of chromatin.

Upon completion of the optimisation steps, a pulse-chase experiment was carried out to ensure that each of the chosen time points would return efficient γ H2AX enrichment. LCLs were irradiated with 20 Gy of IR and they were allowed to recover for a maximum of 24 hours, while taking time points at prescribed times: 5 minutes, 15 minutes, 30 minutes, and 60 minutes. The protein amounts from each sample were normalised and then proceeded to carry out a γ H2AX-IP from the isolated nuclei. Except for T0 which corresponds to mock irradiation, all the time points gave sufficient enrichment of γ H2AX (**Figure 4.11**). The samples were subsequently prepared for analysis via LC-MS. The results for these assays are analysed in Chapter 5.

Figure 4.11 Western blot of γ H2AX IP timecourse.

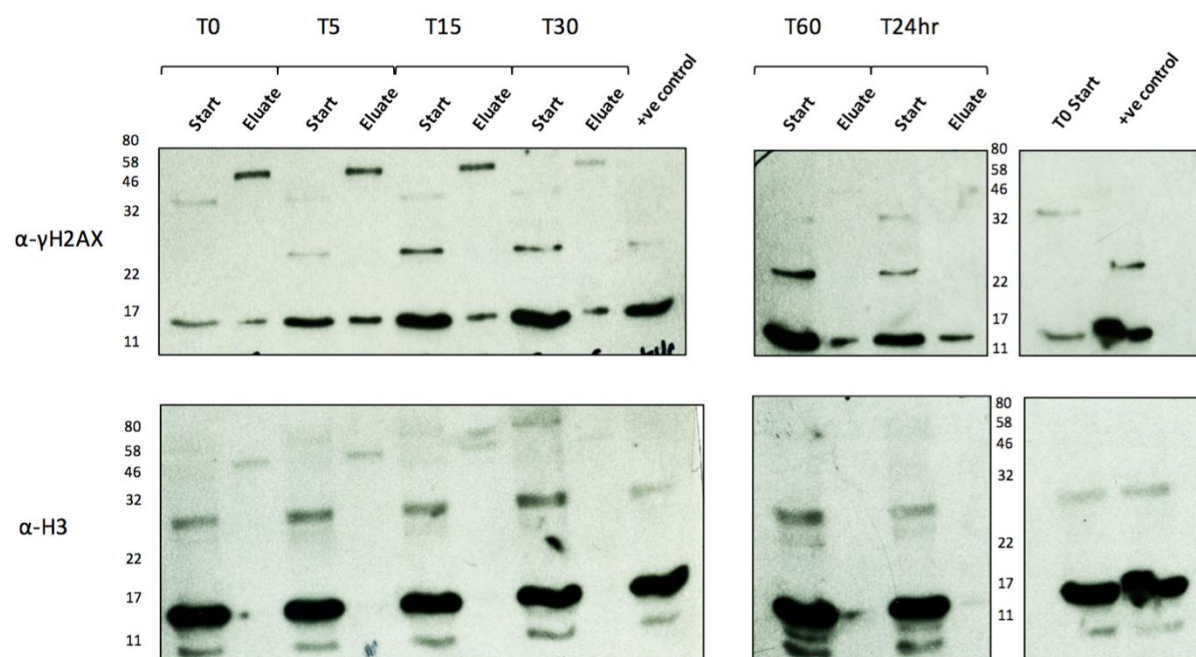


Figure 4.11 Legend: Western blot of γ H2AX IP timecourse.

Western blot of γ H2AX IP timecourse. Pulse chase experiment: LCLs were irradiated with 20 Gy ionizing radiation and were allowed to recover for 5 minutes, 15 minutes, 30 minutes, 60 minutes and 24 hours. 1.5 mg of histone sample was used for each IP sample. Blotted for γ H2AX to measure efficiency of pull down. Histone H3 blotted as a loading control.

Chapter 5: LC-MS Quantification of HPTMs after IR-Induced DNA Damage

5.1 Introduction

As mentioned before, certain HPTMs have been shown to change in response to DSBs, as detected by antibody-based techniques. In Chapter 4, we develop and optimise a novel method for quantifying these changes. A system that should be impervious to load capacity in regards to the number of samples collected and number of different HPTMS assayed simultaneously. It is important to note that even though the readout of the screen for our in-house developed technique is mass spectrometry based, there is still an enrichment step that is dependent on antibodies. The results here will give an indication of whether the method would be viable. Before examining the results, it is crucial to give an outline of what has been reported before in the literature in regards to HPTM changes that occur as a result of the DNA damage response.

5.1.1. H2AX Modifications

γ H2AX is an established indicator of DNA damage and it would be disconcerting if it was not detectable in our system. Any issues in detecting γ H2AX would stem more from a technical issue rather than biological or sample preparation. Unlike typical peptides, the C-terminal tryptic peptide of H2AX has no positively-charged C-terminal residue. Furthermore, phosphorylation makes the peptide more negative. Potential technical issues could arise from analysis in the positive mode. Nevertheless, the interesting point about γ H2AX is not if it is present but how much of it is present. Elucidating the relationship between S139 phosphorylation and Y142 phosphorylation in regards to their

concomitant relationship with apoptosis would also bring much-needed clarity as to what is happening prior to cell death. Since the Allis group reported the phosphorylation of H2AX-Y142 as being ubiquitous and decreasing in the presence of damage (Xiao et al. 2008).

5.1.2. Ubiquitination of H2A(X)

Ubiquitination of H2AX-K13/K15 is synonymous with DNA damage, much like γ H2AX. However, the signal of this ubiquitination is not as intense or quick in formation as the phosphorylation. H2AX-K13/K15 is one of the docking sites for 53BP1, which promotes NHEJ, and since NHEJ is responsible for the repair of most DSBs (Kakarougkas & Jeggo 2014), detecting the ubiquitination should be possible if we pull-down nucleosomes closer to the site of damage. We expect this HPTM to increase and persist at earlier time points but fade away in a pulse-chase experiment. However, detection of this ubiquitination event at a global level may prove too challenging due to the high background associated with equivalent histones distal to sites of DNA damage (i.e., signal from unmodified H2AX-K13/K15 may drown out the modified signal). There have also been reports of ubiquitination at the C-terminal end of the histone tail in response to damage. Classically, Polycomb Repressive Complex 1 is known to ubiquitinate K118/K119 for transcriptional repression. However, Pan et al (Pan et al. 2011) report that a monoubiquitination of H2AX at K120 is required to prime the DDR. If this is the case, we would expect to detect a positive correlation between S139 phosphorylation and K120 ubiquitination.

5.1.3. H3K9 PTMs

H3K9 di- and trimethylation has a role in maintaining the structural integrity of chromatin, as seen in constitutive heterochromatin. That role could serve one of two purposes at

sites of DNA damage: the presence of methylation could stabilise the chromatin so that collateral damage to adjacent chromatin does not take place, or the modification could be removed to allow the repair machinery access to the chromatin. Studies have shown that both can take place: Suv39 and HP1 (human homologues of Swi6 and Clr4, respectively) which are involved in methylating H3K9 and spreading H3K9 methylation, respectively, have both been shown to increase at the site of damage at both euchromatin and heterochromatin (Baldeyron et al. 2011). Ayrapetov et al (2014) suggest that there is a small increase of methylation at the site of damage (Ayrapetov et al. 2014). H3K9Ac was also shown to decrease in response to damage (Tjeertes et al. 2009). We expect that with a pulse-chase experiment, it should be possible to detect if the methylation is decreasing at those points and eventually being restored.

5.1.4. Histone Docking Sites of 53BP1

5.1.4.1. H3K4me2

It has been reported with immune-assays that in response to damage, LSD1 is recruited to the site of DNA damage, resulting in a reduction of H3Kme2 at those sites (Mosammaparast et al. 2013). LSD1 was also shown to be required for recruitment of 53BP1 but not for γ H2AX formation. A link to RNF168 was suggested, which implies that H3K4me2 may be required for repair pathway choice in S and G2 phase. Demethylation of H3K4me2 has been seen to occur during late S and G2. Therefore, when LSD1 is recruited, it may promote HR-mediated DSB repair. H3K4me2 is indicative of transcription and so it could also be that its removal is required to ensure that transcription does not take place until the DNA damage is repaired. We would expect to see an initial decrease in H3K4me2 levels at the site of damage but not at a global level.

5.1.4.2. H4K20me2

Along with H2AK13/K15 ubiquitination, H4K20me1/me2 has also been reported as being required for loading of 53BP1 onto sites of damage (Pellegrino et al. 2017). H4K20me2 is already the most abundant form of H4K20, however docking of 53BP1 onto H4K20me2 is only activated in the presence of DNA damage. If we get enough resolution to distinguish between global H4K20 methylation and local levels, we should see an immediate increase of methylation at local levels. SILAC could also reveal whether any new modifications are taking place or whether a monomethyl is being promoted to dimethylation.

5.1.4.3. H3K79 methylation

H3K79 methylation is present in undamaged cells (Sweet et al. 2010). Nevertheless, it is also required for the recruitment of 53BP1. H3K79 is unique in that it is the only known histone residue that is methylated by a non-SET domain HMT, DOT1. 53BP1 possesses a Tudor domain that will recognise and bind to the H3K79 methylation. 53BP1 is a fast response DNA damage protein, so we expect to see a slight increase compared to global levels soon after the induction of damage.

5.1.5. H3K14ac

H3K14ac had conflicting results in the past in that H3K14ac level has been reported to both increase globally as a result of damage (Kim et al. 2008) but also to exhibit no change in response to damage (Tjeertes et al. 2009). If Kim et al. (2008) are correct, then we should see global increases but decreases at local levels. On the contrary, if Tjeertes et al. (2009) are correct, we should not detect any changes.

5.1.6. H3K36

H3K36me3 is associated with HR-mediated repair of DSBs (Pfister et al. 2014). When the KMT responsible for trimethylating H3K36 was knocked down, there was a significant reduction in the cells undergoing HR. However, if the same residue is demethylated, the DDR's repair pathway choice can be steered towards NHEJ in a Metnase-dependent manner (Fnu et al. 2011). Therefore, what we would anticipate is an increase in both dimethylation and trimethylation of H3K36. However, as mentioned, most of the repair is conducted by NHEJ so we expect to only notice a negligible difference of H3K36me3, whereas the rest of the histone will be disproportionately demethylated.

5.1.7. H3K56ac

In response to damage, H3K56 acetylation has been reported to be required in both yeast and mammalian systems. Tjeertes et al. (2009) detected an immediate decrease of the modification before its levels increased again. The early response of reducing the modification level could be explained as the DDR system requiring to stabilise the chromatin that has been impacted by damage, as H3K56 is known for interacting with the minor groove of DNA (Brehove et al. 2015) and is dependent on H3T45 phosphorylation. Therefore, if there is a sufficient amount of the modification to begin with and if it has good ionization efficiency, we should be able to detect fluctuations in its level by LC-MS/MS after DNA damage.

5.1.8. H4K16Ac

The consensus from previous studies is that H4K16Ac levels will increase at the onset of DNA damage. Deletion of MOF2, the HAT responsible for acetylating H4K16, indicates that it is responsible for ATM recruitment at the site of damage (G. G. Sharma et al. 2010); there were reduced γ H2AX foci, which are a readout for ATM activity, in MOF knock out cells. H4K16ac is also required for dephosphorylation of γ H2AX after DNA damage repair

(Neumayer & Nguyen 2014), so it is feasible that acetylation of H4K16 works as a marker to indicate the progress of the repair to the DDR machinery. Hence, we expect to see increased H4K16ac that persists into the later time points.

5.1.9. H2AZ

While histone H2AZ is not a PTM, it is a variant that has been shown to be involved in the DDR and implicated in several other biological processes: chromatin compaction (Fan et al. 2002), DDR (Y. Xu et al. 2012), meiosis (Yamada et al. 2017), and defining transcriptional start sites and gene bodies (Gervais & Gaudreau 2009) (Coleman-Derr & Zilberman 2012). H2AZ is recruited to sites of DNA damage within mere seconds and it is deposited at such sites. However, within five minutes, it appears to be removed from the site of damage (Gursoy-Yuzugullu et al. 2015) (Alatwi & Downs 2015). If H2AZ is not removed, HR-mediated repair is impeded. H2AZ will be more challenging to quantify as when its DDR-induced chromatin deposit takes place, it would need to be close to a phosphorylated H2AX to be pulled down alongside it.

5.2. Results

5.2.1. Increase in H2AX S139 phosphorylation but not growth in enrichment of other variants of H2A

γ H2AX

Whenever a single or double strand break takes place, H2AX will be phosphorylated. The process is described in much further detail in this thesis in Introduction and Chapter 3. The levels of phosphorylated S139 were measured based on the proportion of ATQASQEY vs. ATQAS[p]QEY normalised to internal control peptides. **(Figure 5.1)**. The difference between the acid extract and Mnase extraction is that the acid extract has the capability to pull down histones from most of the genome indiscriminate of different regions. However, Mnase like any other nuclease will be limited to where it can digest in the genome. Mnase is used in assays to determine if chromatin is more accessible after a given treatment (Mieczkowski et al. 2016).

It is not surprising that enriching for nucleosomes with a α - γ H2AX antibody would result in an increased yield of the γ H2AX signal in the eluate **(Figure 5.1a)**. This control allows us to determine whether the nucleosomes that we enrich for with immuno-based techniques are actually from DNA damaged area. The T0 sample, before induction of damage, still shows increased γ H2AX after IP, consistent with enrichment of background endogenous damage present in the cell. The cells were not synchronised, and γ H2AX occurs at high levels in S-phase **(Figure 5.1b)**.

Another explanation can be attributed to the method development decision of combining the S1 and S2 supernatants after digesting with Mnase **(Figure 4.8)**. S1 would

normally be composed of mono-/di-nucleosomes bound together by the undigested DNA wrapped around their histones and would be more uniform in composition. However, S2 supernatant contains a range in size of digested chromatin from tetranucleosomes to octanucleosomes for example. As seen in **Figure 3.4**, the acid extraction data suggests that approximately 1 in 20 H2A variants are H2AX. Mnase does not digest across the genome uniformly: mainly cutting at euchromatin and we know this is the same region of the genome that H2AX occupies (Hamilton et al. 2014). So there might be a higher proportion of H2AX to start with compared to a crude method of extraction such as acid extraction. However, when we have fragments of chromatin containing upwards of six to ten nucleosomes each and assuming there is a random distribution of H2AX, we would be potentially deflating the H2AX numbers despite enriching for γ H2AX when mixing the S1 and S2 together.

By immunoprecipitating from S1 supernatant solely, the vast majority of the chromatin would have been composed of mono- or-dinucleosomes. The levels of H2AX that we would expect to see would be somewhere approximately of 25-50% depending on the equal distribution of H2AX in the genome and whether H2AX is deposited into the nucleosome symmetrically or asymmetrically i.e. occupies a nucleosome with a different H2A variant.

Figure 5.1. Increase in H2AX S139 phosphorylation but not increase in enrichment of other variants of H2A

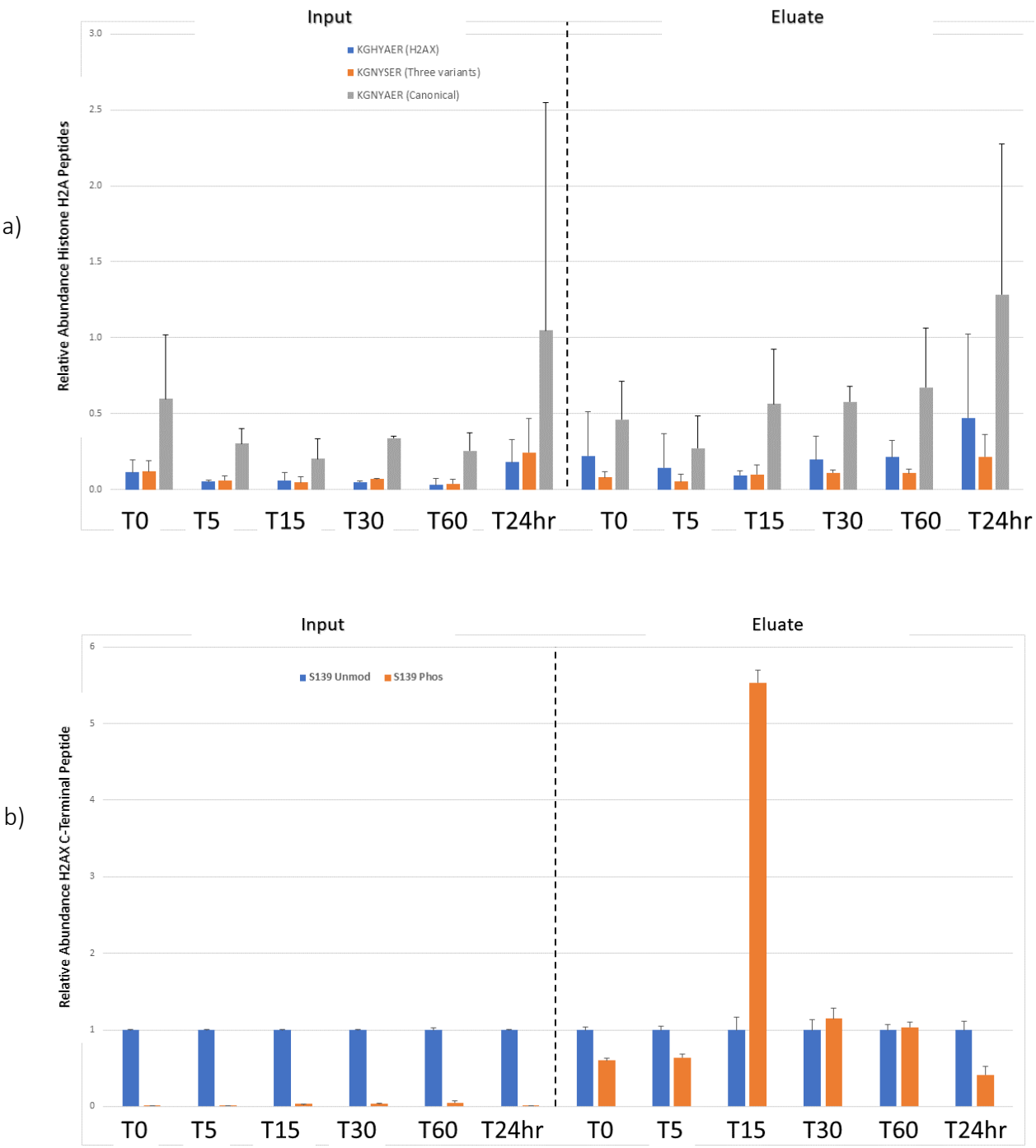


Figure 5.1 Legend: Increase in H2AX S139 phosphorylation but not growth in enrichment of other variants of H2A

- a) Time-course analysis of H2AXS139 post-translational modification changes. LCLs irradiated with 20 Gy ionizing radiation. H2AXS139 modification percentages were quantified by measuring fragment ions for singly charged precursors: ATQASQEY and ATQA[p]SQEY. Error bars correspond to standard deviation. N = 3 biological replicates.

(LHS) There is a steady increase of γ H2AX at a global level that peaks between 30 min and 60 min time point. At 24 hours, γ H2AX levels decrease to that similar of no damage.

(RHS) Enrichment of γ H2AX relative to H2AX. With every time point, there is increased γ H2AX levels especially when compared to the reciprocal global time

- b) Pulse chase analysis of H2AX enrichment post-translational modification changes. LCLs irradiated with 20 Gy ionizing radiation. Percentage of H2AX variants compared to other H2A isoforms. Fragments from doubly charged precursors were used for quantification. Peptides used for quantification: KGHYAER (H2AX), KGNYSER (three variants), and KGNAYER (canonical). Error bars correspond to standard deviation. N = 3 biological replicates. Error bars correspond to standard deviation. N = 3 biological replicates.

(LHS) There is no increase in H2AX at a global level throughout the time points: all ratios remain level.

(RHS) Enrichment of H2AX relative to other H2A isoforms. With larger error bars, there is a trend showing increased H2AX relative to other H2A isoforms. However, even if the increase is taken into consideration, which shows a two-fold enrichment of H2AX relative to the global level, it is not as much as increase as we would expect as we were only pulling out γ H2AX-containing mono-nucleosomes.

5.2.2. Increase in Histone H3K9 Dimethylation

The role that histone H3K9 methylation plays in compacting chromatin has been documented (Introduction & (Saksouk et al. 2015)). In recent years, studies have shown that H3K9 methylation is not only involved in constitutive compaction of areas of the genome void of genes but also has a role to play in the DDR. Baldeyron et al. (2011) show that Heterochromatin Protein 1 α , which co-recruits with the primary HKMT responsible for methylating H3K9 (Suv39) is recruited to the site of damage regardless of if the environment is euchromatic or heterochromatic, in a CAF-1 dependent manner.

The recruitment of HP1 α to the DSBs can take place at both early and late stages of the repair. HP1 α requires KAP1 to be recruited. In a follow-up study from the same group, Soria et al. (2013) find that HP1 α and HP1 β are both involved in DNA end resection which is a prerequisite for homology searching in the homologous recombination repair mediated pathway (Soria & Almouzni 2013). This is in accordance with Goodarzi & Jeggo (2008) (Goodarzi et al. 2008); Sun et al. 2009 (Y. Sun et al. 2009) and Ayrappetov (Ayrappetov et al. 2014). Goodarzi et al. demonstrate that DSB repair in heterochromatic regions requires ATM, to disrupt the binding of KAP1 to heterochromatin by phosphorylating KAP1, especially at later stages of the DNA repair. Sun et al. 2009 showed that Tip60, which activates ATM, requires its chromodomain to bind to H3K9 methylation present at the site of damage for its acetyltransferase activity. In the absence of H3K9 methylation, there is the reduced recruitment of ATM to DSBs. However, H2AX is still phosphorylated by DNA-PKcs. Their data also suggest that there is no change in global levels of H3K9 methylation, therefore in a non-defective environment, Tip60 is recruited to sites of damage where H3K9 methylation is already present i.e., heterochromatin.

However, Ayrapetov et al. (2014) demonstrate that even though there is no global increase in H3K9 methylation levels, there was an increase in levels close to the site of damage. Their ChIP data at a targeted DSB suggested that H3K9me3 levels mirror that of γ H2AX: Increase in H3K9me2/3 levels peak at about 1.5 KB from either side of the DSB and with lower levels about 200 KB. No such effect is seen in undamaged chromatin.

Figure 5.2 shows a significant difference in H3K9me2 levels 24 hours after treatment with IR (with a p-value of 0.02). Goodarzi et al. demonstrate that at later time-points after damage, the percentage of breaks repaired with the NHEJ pathway decreases and skews towards HR-mediated repair. Goodarzi et al. also see that the late-repairing breaks tend to be in a heterochromatic environment. This is a result that was echoed in Natale et al. (2003), where even though they report γ H2AX mainly being present in euchromatin, the signals that persist at later time points were mainly from heterochromatin. This begs the question, are we pulling out γ H2AX from regions of heterochromatin that are repaired late, or does the proportion of histone H3 methylation at the site of damage steadily increase due to chromatin remodelling, i.e., is this pre-existing H3K9 methylation, or damage-induced?

As regards to acetylation of H3K9, multiple studies have shown that there is a marked decrease in a DNA damage environment. In a screen involving immuno-based assays that tracked changes in HPTMs about DNA damage, Tjeertes et al. (2009) observed a decrease in H3K9ac after only fifteen minutes. They treated U2-OS with phleomycin for 2 hours and released the cells back into phleomycin-free medium. Cea et al. (2016) observed that SIRT6, the HDAC responsible for deacetylating H3K9ac, was shown to be crucial for

preventing the proliferation of multiple myeloma cells by destabilising the mitogen-activated protein kinase (MAPK) pathway which promotes cell cycle progression (Cea et al. 2016). SIRT6 binds to the promoter of MAPK and deacetylates H3K9 the transcriptional start site. The knockdown of SIRT6 also resulted in increased sensitivity to DNA-damaging agents such as doxorubicin and melphalan (Cea et al. 2016). Gupta et al. (2016) treated the human malaria parasite *Plasmodium falciparum* with MMS and noticed a similar response akin to that of Tjeertes et al. (2009); a decrease in both H3K9ac and H3K56ac. Tjeertes et al. and Cea et al. both used immune-based assays which will be more sensitive to detecting smaller levels of the acetylated peptides.

Figure 5.2. Increase in Histone H3K9 Dimethylation

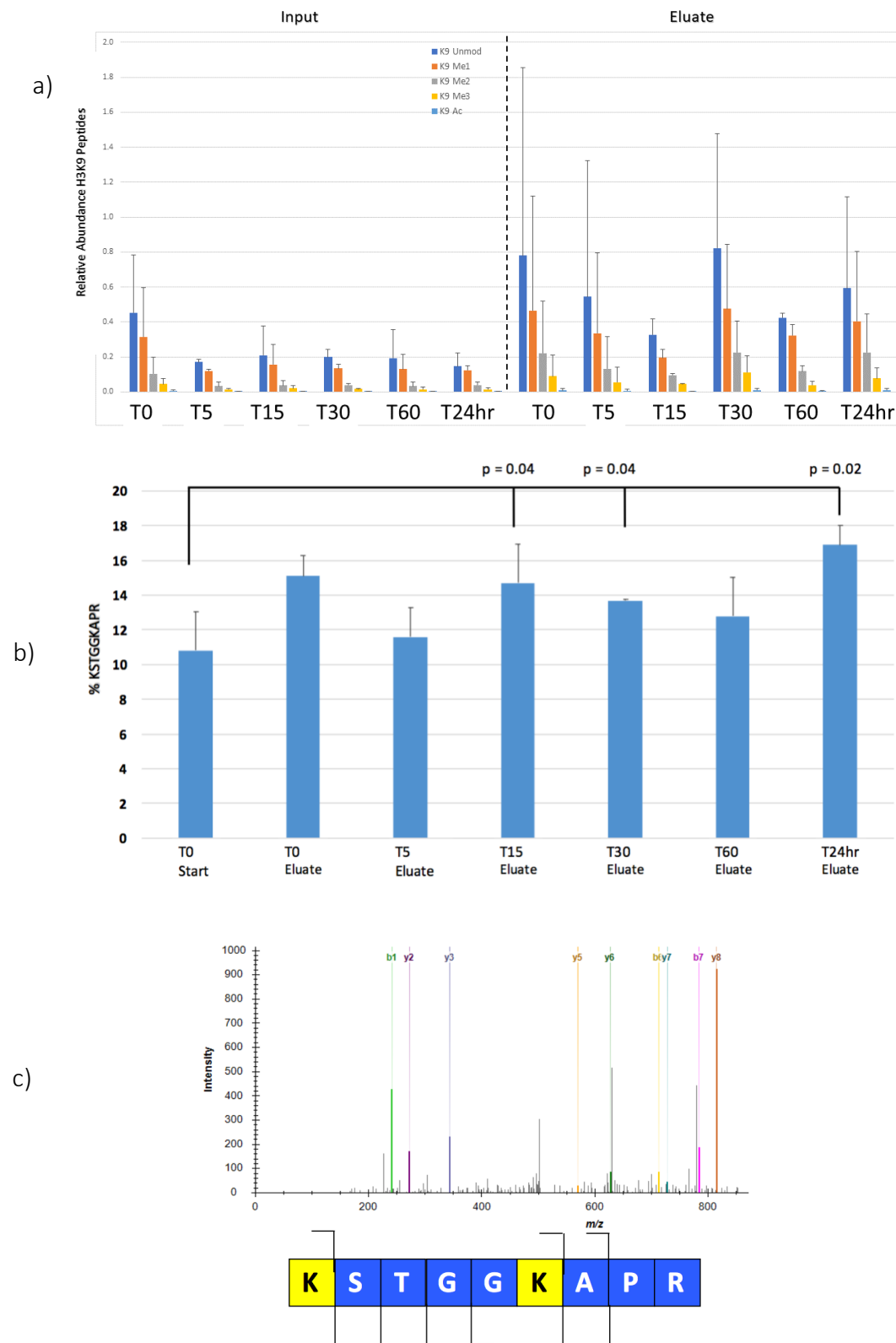


Figure 5.2. Legend: Increase in Histone H3K9 Dimethylation

- a) Time-course analysis of H3K9 post-translational modification changes. LCLs were irradiated with 20 Gy ionizing radiation. H3K9 modification percentages were quantified by measuring fragment ions for doubly-charged precursors. List of peptides used for quantification:
- I. KSTGG[ac]KAPR
 - II. [me1]KSTGG[ac]KAPR,x
 - III. [me2]KSTGG[ac]KAPR
 - IV. [me3]KSTGG[ac]KAPR,
 - V. [ac]KSTGG[ac]KAPR,
 - VI. KSTGGKAPR,
 - VII. [me1]KSTGGKAPR,
 - VIII. [me2]KSTGGKAPR
 - IX. [me3]KSTGGKAPR
 - X. [ac]KSTGGKAPR.
- Error bars correspond to standard deviation. N = 3 biological replicates.
- b) Increased H3K9me2 at the site of damage: 15 mins, 30 mins and 24 hour time points indicate that there is a significant difference in regards to H3K9me2 at the site of damage relative to global average before damage (T0 Start). Two-tailed T-tests were carried out on all time points against undamaged global values.
- c) (LHS) MS/MS fragmentation of doubly charged precursors for residues corresponding to modified KSTGGKAPR peptide. Fragment ions involved in quantitation are highlighted on the peptide sequence: b1, b6, b7, y2, y3, y5, y6, y7, and y8.

5.2.3 There is no difference in Histone H3K27 methylation levels post DNA damage

Histone H3K27 methylation is another repressive mark of chromatin. In the native cell environment in the absence of damage, H3K9 methylation is found at highly repetitive sequences such as at telomeric and pericentromeric sites so the location in the genome at which H3K9 is expected to occupy would be entirely consistent with one cell type to another (Lomber et al. 2006). In comparison, from one cell to another, H3K27 methylation occupies different genes due to cell differentiation contextually requiring silencing of genes (Beisel & Paro 2011).

The literature suggests that there should be an increase in H3K27 methylation, especially the trimethyl modification (Campbell et al. 2014). At the site of damage, PRC2 will trimethylate H3K27. H3K27me3 recruits PRC1 which will ubiquitinate H2AK119 in a PARP-dependent manner (H. Wang et al. 2004; Campbell et al. 2014). H3K27me3 can recruit more PRC1 and PRC2 in both replication and repair which in turn methylates H3K27 on adjacent nucleosomes. Therefore it is a self-propagating modification (H. Wang et al. 2004). Campbell et al. (2014) notice that the recruitment of PRC1 and PRC2 to the site of damage occurs in an H2AX/PIKK kinases independent manner.

MacroH2A.1 can be found in two regions of the genome in equal proportions: sites that are hyperacetylated (H2AK5, H2BK12, K15, K20, K120, H3K4, H3K14, K18, and H4K91) indicative of expressive sites; and H3K27me3 which as mentioned is a site of facultative heterochromatin (H. Chen et al. 2014). The region of the genome that would contain nucleosomes H3K27me3 and MacroH2A.1 can span 500 KB in length (Gamble et al. 2010). The two domains strictly do not overlap. This is critical in regards to our H3K27 data for

two reasons (**Figure 5.3**): if we are enriching nucleosomes with a γ H2AX antibody, then it's less likely that we are going to be pulling out MacroH2A.1 as they would not necessarily be occupying the same nucleosomes. Also, γ H2AX after ionising radiation are mainly found in euchromatic regions rather than heterochromatic regions (Cowell et al. 2007). Inevitably, enrichment of H3K27me3 may be reduced from the onset.

Figure 5.3: No difference in Histone H3K27 methylation levels post DNA damage

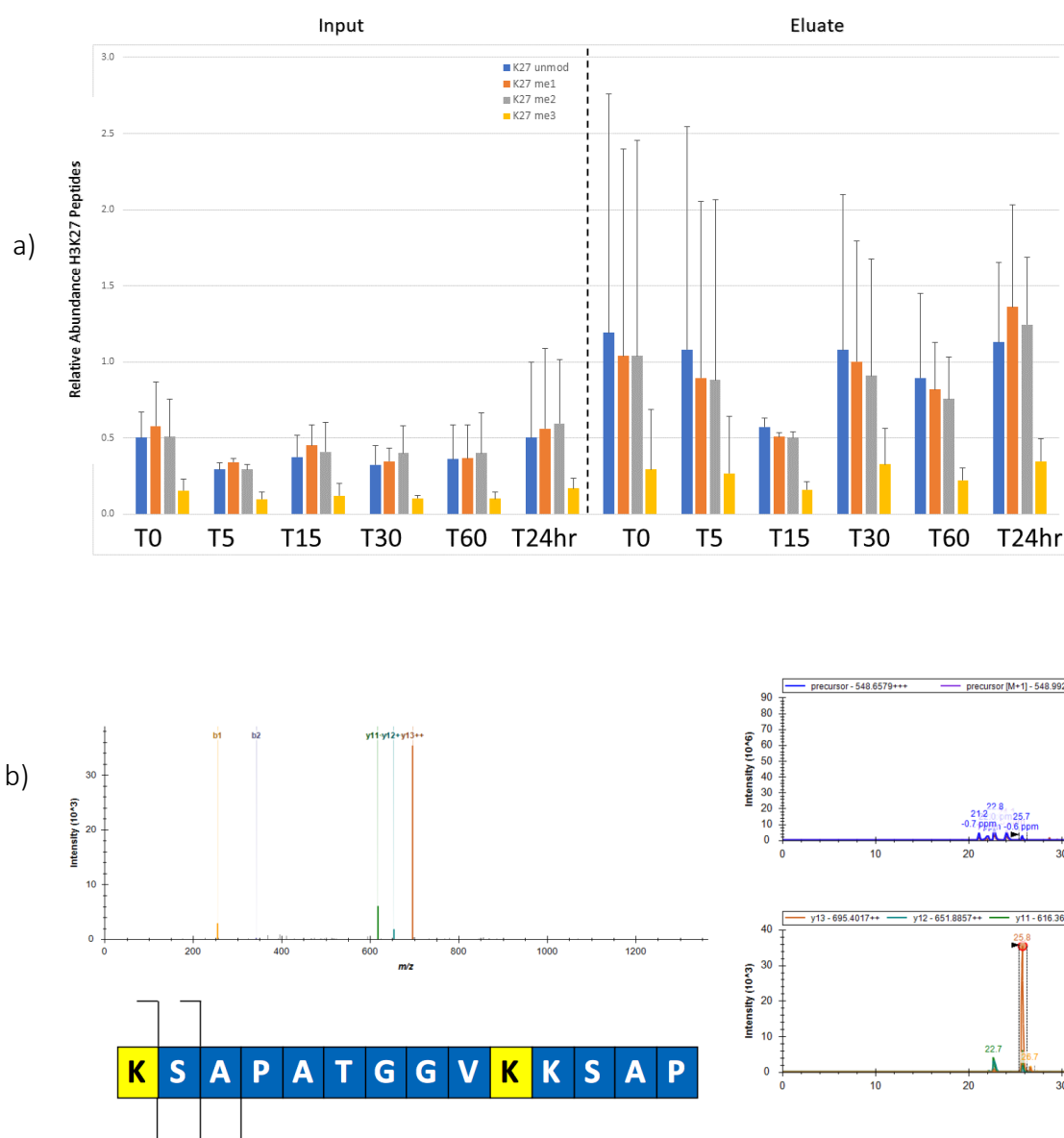


Figure 5.3: No difference in Histone H3K27 methylation levels post DNA damage

a) Pulse chase analysis of H3K27 post-translational modification changes. LCLs irradiated with 20 Gy ionizing radiation. H3K27 modification percentages were quantified by measuring fragment ions for triply charged precursors. List of peptides used for quantification:

- I. KSAPATGGVKK SAP
- II. KSAPATGGV[me1]KK SAP
- III. KSAPATGGV[me2]KK SAP
- IV. KSAPATGGV[me3]KK SAP
- V. [me1]KSAPATGGVKK SAP
- VI. [me1]KSAPATGGV[me1]KK SAP
- VII. [me1]KSAPATGGV[me2]KK SAP
- VIII. [me1]KSAPATGGV[me3]KK SAP
- IX. [me2]KSAPATGGVKK SAP
- X. [me2]KSAPATGGV[me1]KK SAP
- XI. [me2]KSAPATGGV[me2]KK SAP
- XII. [me2]KSAPATGGV[me3]KK SAP
- XIII. [me3]KSAPATGGVKK SAP
- XIV. [me3]KSAPATGGV[me1]KK SAP
- XV. [me3]KSAPATGGV[me2]KK SAP

Error bars correspond to standard deviation. N = 3 biological replicates.

b) (LHS) MS/MS fragmentation of triple charged precursors for residues corresponding to modified KSAPATGGVKK SAP peptide. Fragment ions involved in quantitation are highlighted on the peptide sequence: b1, b2, y11, y12, and y13. (RHS) Intensity of weakest sample analysed with pseudo-SRM MS/MS

5.2.4. There is no difference in Histone H3K36 di-/trimethylation levels post DNA damage

A difference of one methylation on histone H3K36 can reportedly result in the DDR being guided towards a different repair pathway. In mammalian systems, histone H3K36me2 recruits NHEJ repair proteins in the presence of a break. Fnu et al. (2011) demonstrated that at the site of damage, H3K36 is dimethylated in a Metnase dependent manner. After the induction of DSBs, they did not see an increase in the trimethylated form. However, after inducing a site-specific break using I-SceI endonuclease followed by ChIP of both metnase and H3K36me2, they saw an overlapping enrichment of the two (neither were present before damage). Overexpression of Metnase resulted in increased enrichment of Ku70, DNA LigIV and phosphorylated NBS1 which are proteins involved in the NHEJ pathway (Fnu et al. 2011).

H3K36 is trimethylated by the HMT SETD2/HYPB. When H3K36 possesses a triple methylated modification, it can push the repair pathway choice towards HR-mediated repair. Pfister et al. (2010) demonstrated that by either knocking down SETD2, overexpressing KDM4A (KDMT for H3K36me3) or by introducing H3.3 that was mutated at Lysine 36 with a methionine would all lead to the same result: reduced recruitment of Rad51 and RPA to the DSB. LEDGEF binds to H3K36me3 with its PWWP domain, which in turn recruits CtIP. CtIP instigates DSB end resection which is a requirement for HR-mediated repair to take place. (Pfister et al. 2014). Pfister et al. did not see a change in levels of H3K36me3, even at the DSB. Fnu et al. (2011) did not notice any change in H3K36me3 local levels either, but they did see a sharp increase in H3K36me2 which would steer the repair pathway towards NHEJ.

Cao et al. (2015) (Cao et al. 2015) show that at the site of damage and in the presence of H3K36me2, ATM phosphorylates KDM2. Once phosphorylated KDM2 loses its chromatin binding ability, and as a consequence H3K36me2 levels increase at the site of damage. JMJD-5 KDM8 is shown to be required to keep levels of H3K36me2 in check for HR-mediated late repairs of breaks to take place unimpeded in *C. elegans* (Amendola et al. 2017). From our data (**Figure 5.4**) we detect no change in H3K36me3; this is in agreement with both Pfister et al. and Fnu et al. However, we did expect to see an increase in H3K36me2.

Figure 5.4 No difference in Histone H3K36 di-/trimethylation post DNA damage

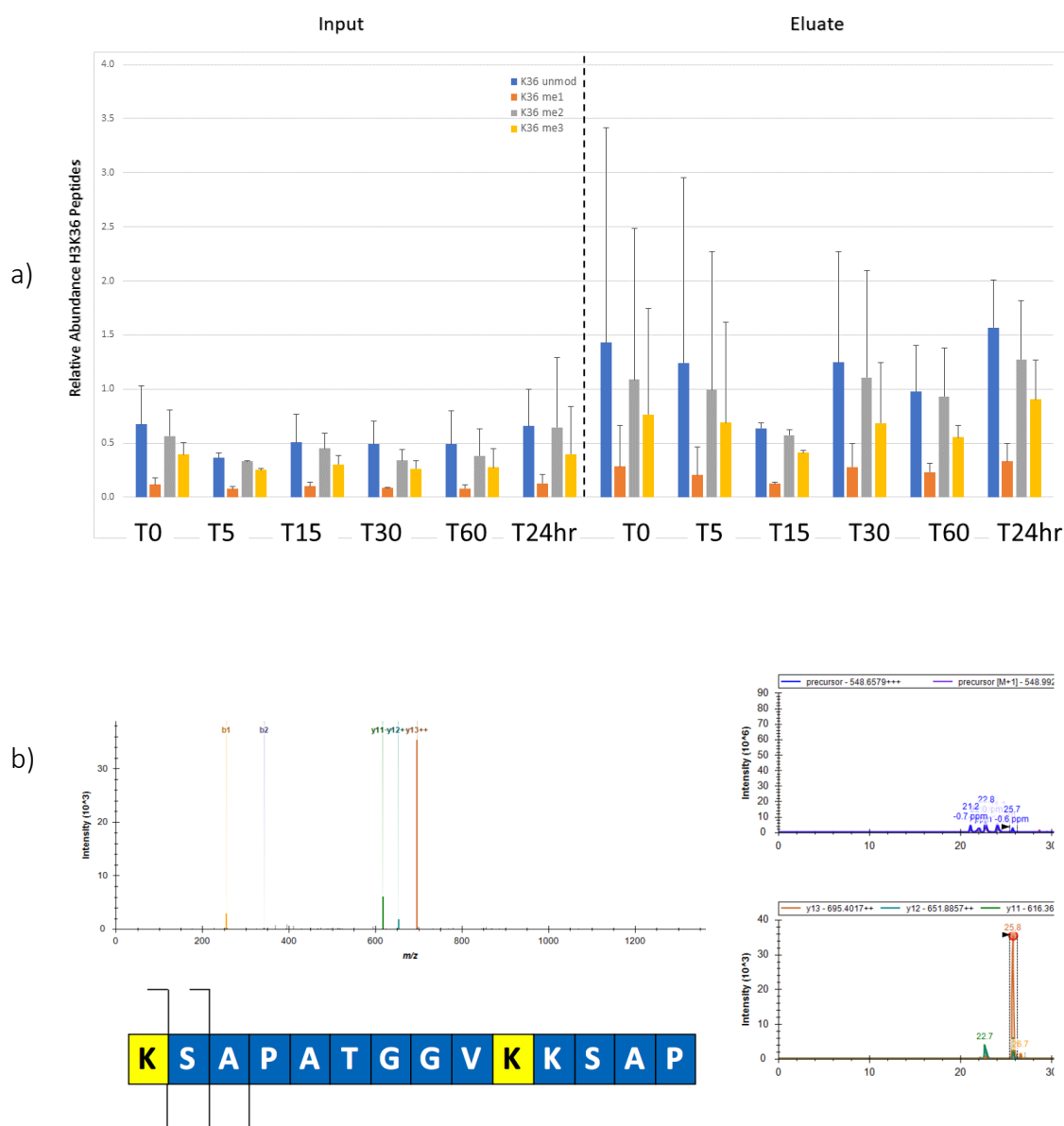


Figure 5.4 Legend: No difference in Histone H3K36 di-/trimethylation post DNA damage

a) Pulse chase analysis of H3K27 post-translational modification changes. LCLs irradiated with 20 Gy ionizing radiation. H3K27 modification percentages were quantified by measuring fragment ions for triply charged precursors. List of peptides used for quantification:

- I. KSAPATGGVKK SAP
- II. KSAPATGGV[me1]KKSAP
- III. KSAPATGGV[me2]KKSAP
- IV. KSAPATGGV[me3]KKSAP
- V. [me1]KSAPATGGVKK SAP
- VI. [me1]KSAPATGGV[me1]KKSAP
- VII. [me1]KSAPATGGV[me2]KKSAP
- VIII. [me1]KSAPATGGV[me3]KKSAP
- IX. [me2]KSAPATGGVKK SAP
- X. [me2]KSAPATGGV[me1]KKSAP
- XI. [me2]KSAPATGGV[me2]KKSAP
- XII. [me2]KSAPATGGV[me3]KKSAP
- XIII. [me3]KSAPATGGVKK SAP
- XIV. [me3]KSAPATGGV[me1]KKSAP
- XV. [me3]KSAPATGGV[me2]KKSAP

Error bars correspond to standard deviation. N = 3 biological replicates.

b) (LHS) MS/MS fragmentation of triply charged precursors for residues corresponding to modified KSAPATGGVKK SAP peptide. Fragment ions involved in quantitation are highlighted on the peptide sequence: b1, b2, y11, y12, and y13. (RHS) Intensity of weakest sample analysed with pseudo-SRM MS/MS

5.2.5. There is a decrease in H3K79me1 at sites local to damage

Histone H3K79 is unique in that it is the only histone residue that is methylated by a non-SET domain methyltransferase, the HKMT DOT1L. Generally, levels of methylation through the cell cycle tend to be low. Levels for the me0, me1, me2 and me3 are around 88%, 5-11%, 1.5% and <0.1% respectively (Sweet et al. 2010). The MS/MS signal for the trimethylated form of H3K79 was below the levels of detection. Previous studies showed that H3K79me2 levels are low during S-phase but become elevated upon DSB (Wakeman et al. 2012) (Sweet et al. 2010). H3K79me2, along with several other residues, are required for the docking of 53BP1 to the site of damage. DOT1 requires Bat3 to co-localise at Histone H3. If either one of Bat3 or Dot1 is knocked down, there is reduced 53BP1 foci at DSBs; even though 53BP1 has a higher affinity for H4K20me2 (Wakeman et al. 2012). In addition, H2BK120Ubi was also shown to be crucial for the DOT1L to be able to bind to H3. H2BK120Ubi alters the stoichiometry between DOT1L and the H3K79me binding site, resulting in increased affinity. (Zhou et al. 2016).

However, it is difficult to ascertain whether our result of decreased monomethylation is due to the H3K79me1 being demethylated or methylated further (**Figure 5.5**). The least convoluted explanation would be that H3K79 methylation plays a critical role in constitutive heterochromatin formation and as mentioned γ H2AX is mainly found in areas of the genome that are more transcriptionally active (Cowell et al. 2007). However, this interpretation is difficult to reconcile with the unchanged global and γ H2AX-associated K9me2,me3 and K27me3 profiles we record. To control for the possible lack of overlap between γ H2AX and H3K79 methylation we would need to use a α -H2AX antibody when enriching for nucleosomes from cells that have been treated with IR. The

idea being that H2AX should be more evenly distributed throughout the genome compared to the phosphorylated form.

Another possible interpretation for the decrease in H3K79me1 levels is that the H3K79me1 are being “promoted” to methylated rank. The reasoning is because Dot1L’s method of action in generating H3K79me2 is through methylating unmethylated or monomethylated H3K79 residues. Therefore, if H3K79me1 is being methylated further then there would be a decrease in monomethylation. The issue with this interpretation is that we do not have data to suggest that there is a change in the H3K79me2. There are no known demethylases for K79.

Figure 5.5. Decrease in H3K79me1 at sites local to damage

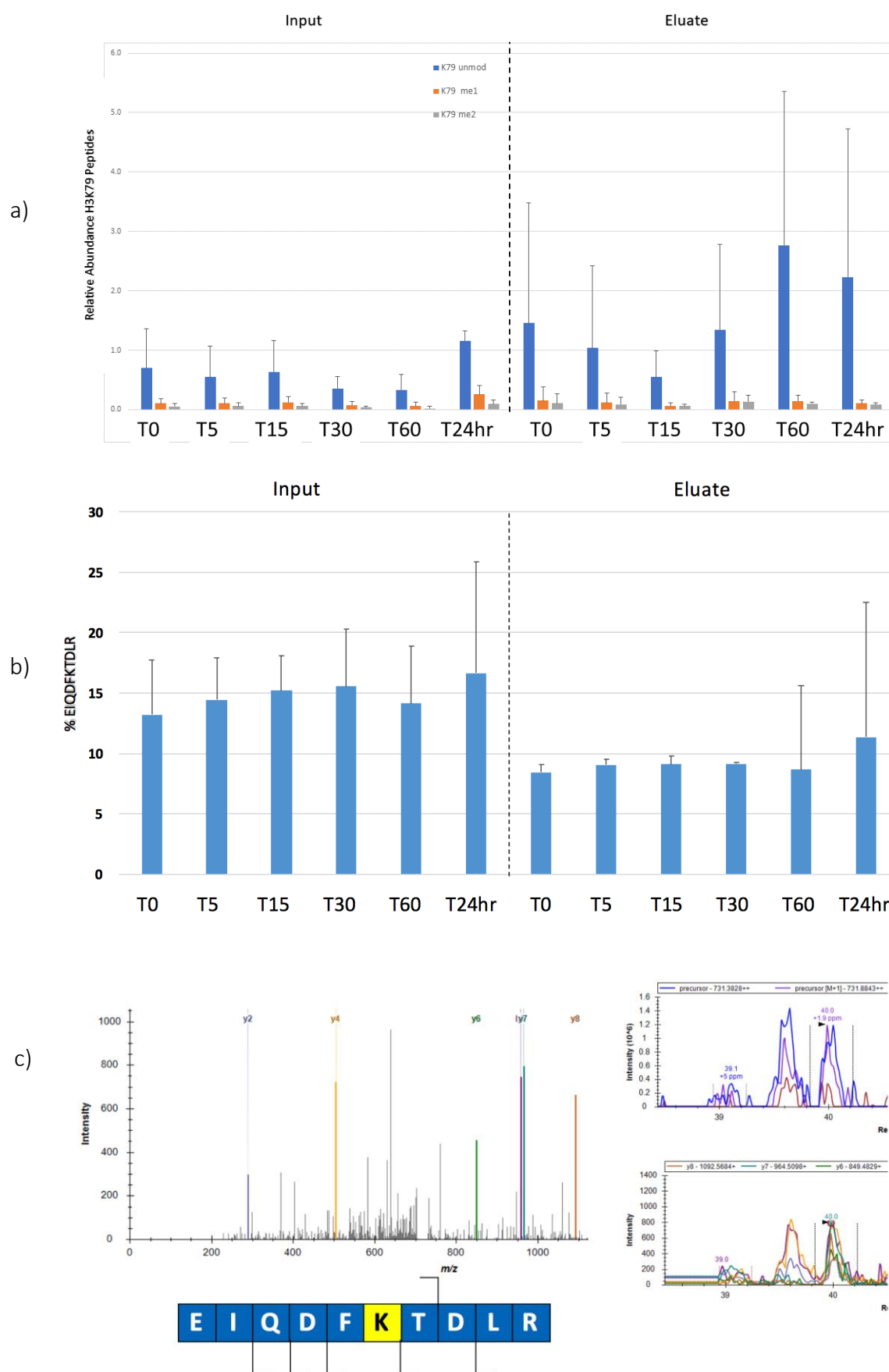


Figure 5.5. Legend: Decrease in H3K79me1 at sites local to damage

- a) Pulse chase analysis of H3K79 post-translational modification changes. LCLs irradiated with 20 Gy ionizing radiation. H3K79 modification percentages were quantified by measuring fragment ions for doubly charged precursors. List of peptides used for quantification:
 - a. EIQDFKTDLR
 - b. EIQDF[me1]KTDLR
 - c. EIQDF[me2]KTDLR

Error bars correspond to standard deviation. N = 3 biological replicates.

- b) Decreased H3K79me1 at the site of damage: 5 mins, 15 mins, and 30 mins time points indicate that there is a significant difference in regards to H3K79me1 at the site of damage relative to global average before damage.
- c) (LHS) MS/MS fragmentation of doubly charged precursors for residues corresponding to modified EIQDFKTDLR peptide. Fragment ions involved in quantitation are highlighted on the peptide sequence: b8, y2, y4, y6, y7, and y8. (RHS) Intensity of weakest sample analysed with pseudo-SRM MS/MS

5.2.6. There is a decrease in H4K20me1 at local sites of damage

Methylation of histone H4K20 has been implicated in several biological processes. Similar to H3K9 methylation, H4K20 trimethylation is a hallmark of constitutive heterochromatin. Suv420 is a histone methyltransferase responsible for spreading H4K20 methylations marks downstream of Suv39 (Jones et al. 2008). Therefore, H4K20 methylation will be found in abundance at pericentromeric and telomeric regions. 53BP1 can be promiscuous in where it docks on the nucleosome after a DSB. However, 53BP1 has a very high affinity for H4K20me2 via its Tudor domain (Panier & Boulton, 2014). H4K20me2 levels are at 85% within the genome (Pesavento et al. 2007) which then begs the question of why does 53BP1 not continuously bind to H4K20me2 outside of the context of DSBs. L3MBTL1 and KDM4A will compete with 53BP1 until damage recognition, where RNF168 is responsible for deactivating both L3MBTL1 and KDM4A binding to H4K20me2. In turn, allowing 53BP1 to bind to the nucleosome in the presence of DSBs (Mallette et al. 2012). Additionally, 53BP1 binding requires the damage-dependent RNF169 ubiquitination of H2A K13/K15 (Fradet-Turcotte et al. 2013).

A decrease in H4K20me1 was observed similar to what is seen with H3K79me1 levels decreasing at local levels, there are various explanations (**Figure 5.6**). H4K20me1 is getting methylated further towards H4K20me2. H4K20 methylation is also found more in constitutive heterochromatin (Cowell et al. 2007) so it is possible that there would be a decreased overlap in the genomic regions where H4K20me2 and γ H2AX also occupy. H4K20me1 has been shown to associate with H3K9me3 in a repressive manner (Barski et al. 2007) This could be tested with a α -H2AX antibody similar to H3K79. It has been suggested that MMSET can methylate H4K20. In response to DNA damage, there is

increased localised methylation at the site of DSB even though there is not an overall growth in global levels of H4K20 methylation. If MMSET is downregulated there is a significant decrease of H4K20 methylation leading to declining of 53BP1 recruitment at the site of damage (Pei et al. 2011b).

At DSBs induced with the I-SceI break site, Pei et al. (2011) observe an increase of H4K20me1/me2/me3. When they used IR to induce DNA damage at lower levels, they detect an increase in H4K20me2 signal. However, when they used lower levels of IR they did not see an increase in H4K20me2 local levels with Western blotting; the implication was that H4K20me2 increases due to low levels of IR were being concealed by the high levels of H4K20me2 within the genome. In this assay, the LCLs were irradiated with a large dose of radiation, it is possible that the high levels of H4K20me2 present within the genome already present would be obscuring any changes that took place due to damage.

Figure 5.6. Decrease in H4K20me1 at local sites of damage in irradiated LCLs

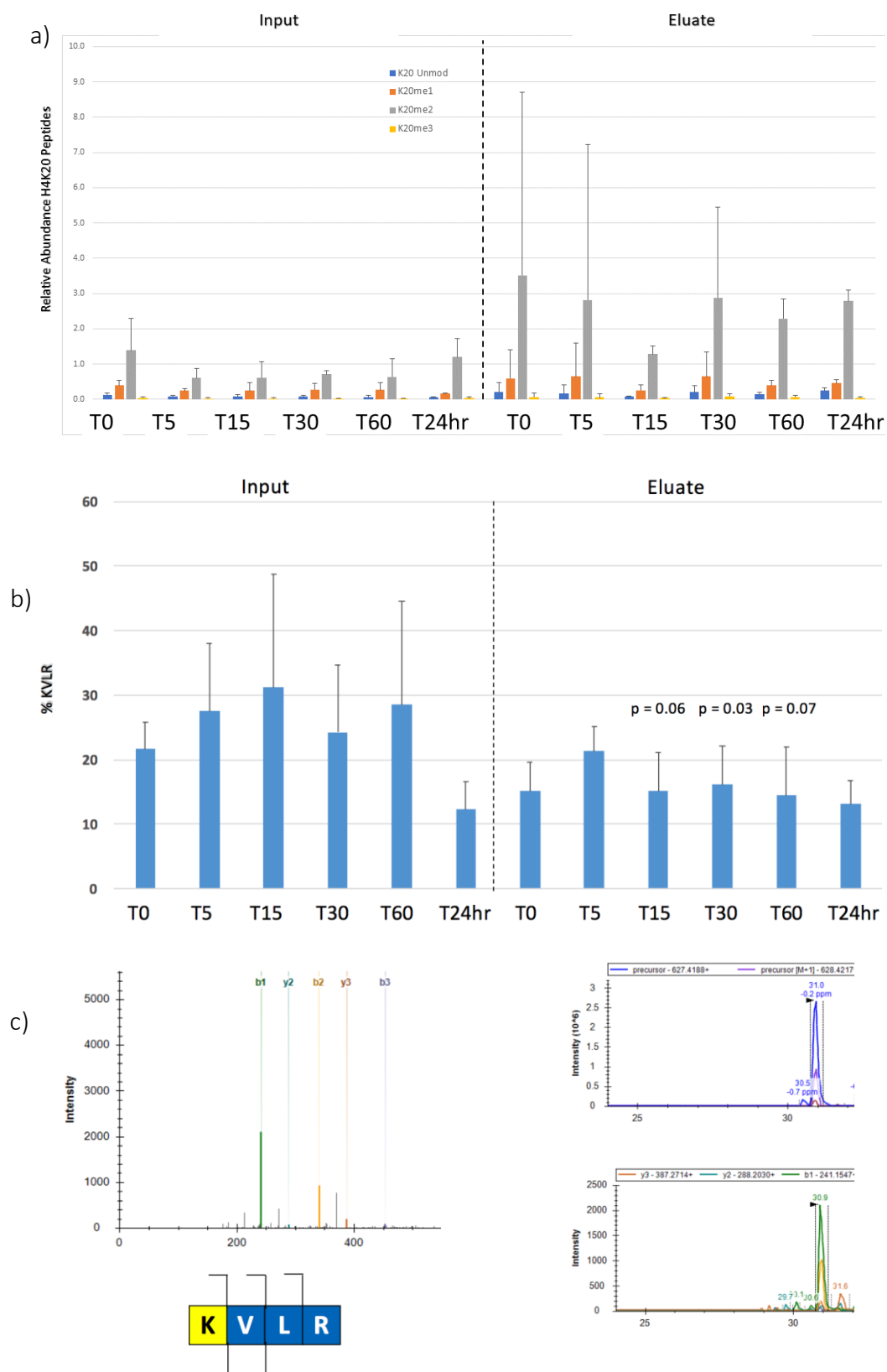


Figure 5.6. Legend: Decrease in H4K20me1 at local sites of damage in irradiated LCLs

- a) Pulse chase analysis of H4K20 post-translational modification changes. LCLs irradiated with 20 Gy ionizing radiation. H420 modification percentages were quantified by measuring fragment ions for singly charged precursors. List of peptides used for quantification:
- a. KVLRL
 - b. [me1]KVLRL
 - c. [me2]KVLRL
 - d. [me3]KVLRL

Error bars correspond to standard deviation. N = 3 biological replicates.

- b) Decreased H4K20me1 at the site of damage relative to global: 15 mins, 30 mins, and 60 mins time points indicate that there is a trend in regards to decreased H4K20me1 at the site of damage relative to global average before damage.
- c) (LHS) MS/MS fragmentation of singly charged precursors for residues corresponding to modified KVLRL peptide. Fragment ions involved in quantitation are highlighted on the peptide sequence: b1, b2, b3, y2, and y3. (RHS) Intensity of weakest sample analysed with pseudo-SRM MS/MS

5.2.7. Small increase in Histone H4 Lysine acetylation post DNA damage

H4K16 acetylation is an HPTM that plays a crucial role in the DDR. MOF is the HAT responsible for acetylating the vast majority of H4K16 within the genome (G. G. Sharma et al. 2010). Even though it is an acetylated modification, H4K16ac is not associated with active transcription. When MOF was knocked down, there were reduced levels of γ H2AX. In addition, the formation of γ H2AX foci were retarded (G. G. Sharma et al. 2010). They show that the HAT activity of MOF was also required for inducing ATM activity (G. G. Sharma et al. 2010). Mice cells exhibited reduced viability when MOF had reduced association with chromatin in Zmpste24-deficient mice (Krishnan et al. 2011).

The increase in H4 lysine acetylation that was observed (**Figure 5.7a**) is possibly due to ATM being required at late repairing breaks, so there is a spike in MOF activity later on in the time points. However, it is not a significant difference as the error bars are overlapping. H4K16ac also plays a role in pathway choice. As mentioned, late repairing breaks undergo repair through the HR-mediated pathway in an ATM-KAP1-dependent manner. However, apart from the 24-hour time point, there is no noticeable difference, and the data are quite noisy. This result is contradictory to what has been seen in previous studies where H4K16ac is required in the DDR: Sharma et al. (2010) irradiated 293T cells with 6 Gy of IR. Two hours post-IR treatment, there was an increased association of MOF with DSBs.

Another role that H4K16ac has in the DDR is once repair of the DSB has taken place, it is responsible for initiating dephosphorylation of γ H2AX (Neumayer & Nguyen 2014). TPX2 plays a role in controlling γ H2AX levels post DNA damage, and depletion of TPX2 will

result in decreases in acetylation of H4K16. Knockdown of SIRT1 (HDAC of H4K16ac) results in increased H4K16ac leading to decreased γ H2AX. Therefore, it is possible that at the site decreased levels of H4K16ac would be detected. However, these findings go against a magnitude of studies of what has been reported prior to Sharma et al. and other studies. Krishnan et al. (2012) also show that H4K16ac depletion could result in early cell death. It is possible to conceive that the cells were very sick as a result of being irradiated with 20Gy of gamma irradiation. The high ionising radiation may have induced apoptosis as their repair machinery were overwhelmed with the amount of radiation that the cells were subjected to 2Gy of ionising radiation in humans is lethal. Another reason to be sceptical of the data is that the H4K16 acetylation is a modification involved in the DDR that is conserved from mammalian to malaria (*P. falciparum*) (Gupta et al. 2016).

HAT1, responsible for acetylating H4K5ac, is involved in mammalian DNA damage repair and is commonly overexpressed in cancer cell lines (Yang et al. 2013). HAT1 depletion impedes HR-mediated repair but not NHEJ repair. Not detecting any changes H4K5ac levels post-DNA damage (**Figure 5.7b**) goes against what was seen in other studies, which further increased our scepticism with the H4 acetylation.

Total acetylation levels across four H4 lysine residues post DNA damage were measured (**Figure 5.7c**). Histone H4 has a rich lysine cluster at N-terminal tail that has been reported to have increased signal in lysine acetylation post-DSB. Several histone acetyltransferases can acetylate the residues. The histone H4 residues H4K5, K8, K12 and K16, were quantified for acetylation simultaneously i.e. four acetylation modifications on histone H4 tail. H4 acetylation increases have been reported in numerous studies across different

systems. The other factor for that being an expected result is that γ H2AX tends to locate around euchromatin and lysine acetylation is normally indicative of more open chromatin environment. In yeast, the HAT1 homologue, ESA1, was shown to be required for efficient repair of DSBs. In *esa1* mutants, the cells have increased sensitivity to camptothecin, and the Histone H4 tails are hypoacetylated (Bird et al. 2002). In *gcn5* mutants, the cells were also shown to be sensitive to constant exposure of the HO endonuclease and cleavage. In ChIP experiments, Gcn5 would be found to enrich proximally to the break site. They found increased levels of acetylation 0.6KB away from the HO site. At 2KB from the DSB was not as much of an increase in acetylation. The role of HATS in the DDR is conserved from yeast to humans. HATS have been shown to play a critical role in HR-mediated repair (Yang et al. 2013).

We finally quantified the enrichment of unacetylated histone H4 across the genome and local to the site of damage (**Figure 5.7d**). What we see is a marked decrease of non-acetylated histone H4 lysines in the 24-hour time point, which is not a surprise as it mirrors what was seen before. Late repairing breaks may give an indication as to why there is a decrease in HK16 deacetylation as those breaks need to be repaired via HR-mediated repair (Tamburini & Tyler 2005).

Figure 5.7. Small increase in Histone H4 Lysine acetylation post DNA damage

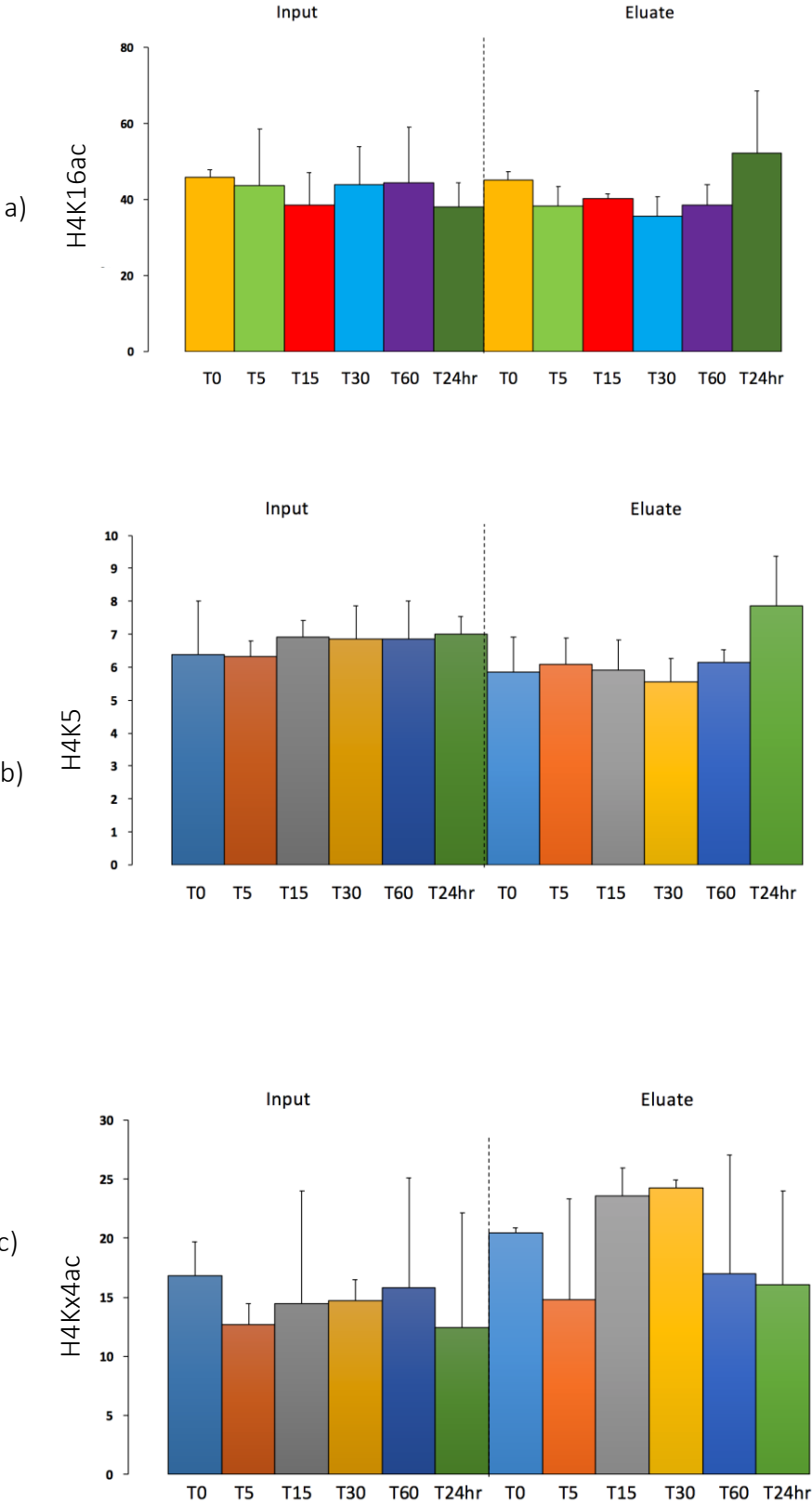


Figure 5.7. Legend: Small increase in Histone H4 Lysine acetylation post DNA damage

- a) Time-course of H4K16ac post-translational modification changes. LCLs irradiated with 20 Gy ionizing radiation. H4K16ac modification percentages were quantified by measuring fragment ions for doubly charged precursors. List of peptides used for quantification: KSTGG[ac]KAPR, [me1]KSTGG[ac]KAPR, [me2]KSTGG[ac]KAPR, [me3]KSTGG[ac]KAPR, [ac]KSTGG[ac]KAPR, KSTGGKAPR, [me1]KSTGGKAPR, [me2]KSTGGKAPR, [me3]KSTGGKAPR, and [ac]KSTGGKAPR. Error bars correspond to standard deviation. N = 3 biological replicates.
- b) Pulse chase analysis of H4K5 post-translational modification changes. LCLs irradiated with 20 Gy ionizing radiation. H4K5ac modification percentages were quantified by measuring fragment ions for singly and doubly charged precursors. Error bars correspond to standard deviation. N = 3 biological replicates.
- c) Pulse chase analysis of H4 quadruple acetylation changes. LCLs irradiated with 20 Gy ionizing radiation. H4Kx4ac modification percentages were quantified by measuring fragment ions for singly and doubly charged precursors. Error bars correspond to standard deviation. N = 3 biological replicates.

Figure 5.7. Small increase in Histone H4 Lysine acetylation post DNA damage
(Continued)

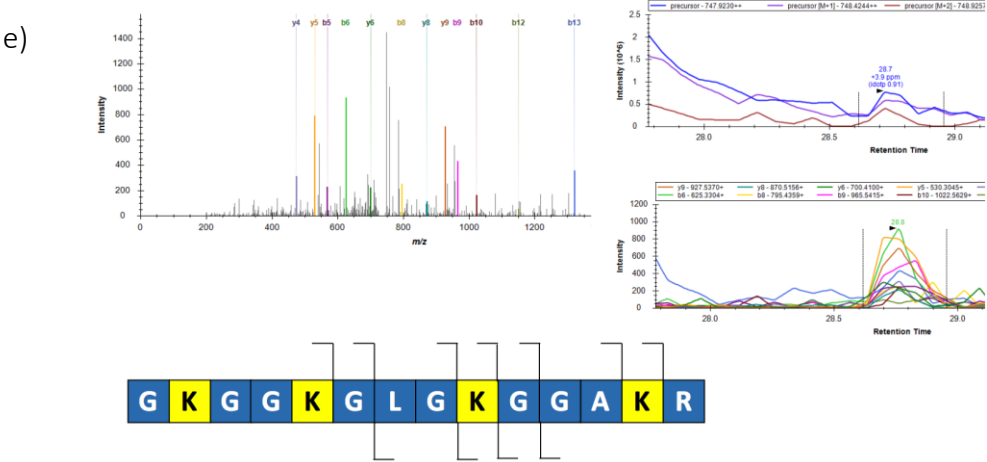
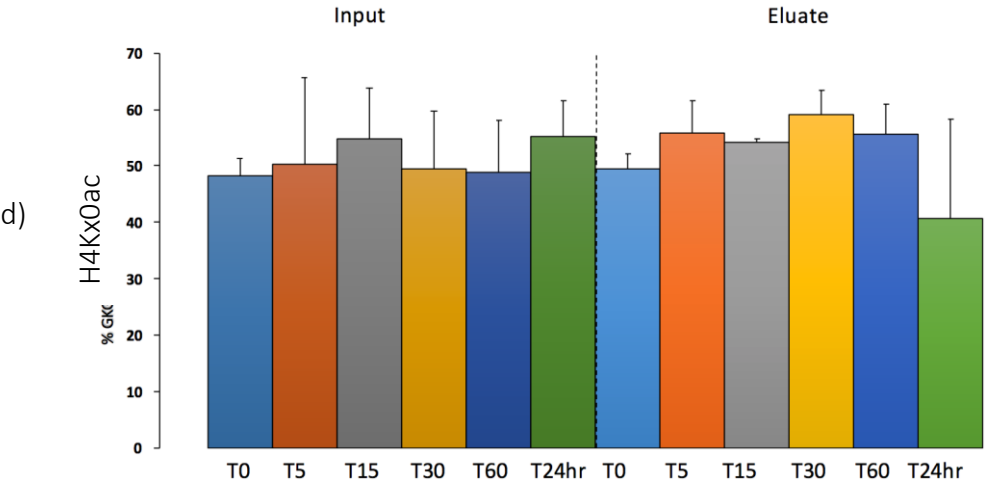


Figure 5.7. Legend: Small increase in Histone H4 Lysine acetylation post DNA damage

- d) Pulse chase analysis of H4 no-acetylation changes : LCLs irradiated with 20 Gy ionizing radiation. H4Kx0ac modification percentages were quantified by measuring fragment ions for singly and doubly charged precursors. Error bars correspond to standard deviation. N = 3 biological replicates.
- e) (LHS) MS/MS fragmentation of doubly charged precursors for residues corresponding to modified GKGGKGLGKGGAKR peptide. Fragment ions involved in quantitation are highlighted on the peptide sequence: b5, b6, b8, b9, b10, b12, b13, y4, y5, y6, and y8. (RHS) Intensity of weakest sample analysed with pseudo-SRM MS/MS

5.2.8. H3K14 acetylation does not change post-DNA damage at either local or global levels

Histone H3K14 acetylation has previously been reported to increase, decrease or not decrease at all where changes in HPTM levels as a consequence of DNA damage were assessed. Both Yang (2013) and Tamburini (2005) reported an increase in the I-SceI and HO cutting sites respectively. Kim (2008) reported a global increase of H3K14 acetylation but a local decrease following IR treatment. Tjeertes et al. (2009) on the other hand did not see any reproducible change following induction of damage with either of camptothecin, phleomycin, IR, MMS, UV, or H₂O₂ in U2-OS cell line.

The H3K14 analysis (**Figure 5.8**) produces some of the most convincing data amongst the HPTMs that were assayed because the error bars are much tighter: adding a significant degree of certainty that H3K14 acetylations do not change following IR damage at the least since no other other damaging agents were used. No changes were detected at neither global levels or local levels and neither at early time points or late time points. 25% of H3K14 is acetylated, regardless of enrichment of at local levels of γ H2AX or global levels. When measured, global H3K14 acetylation levels in U2-OS cells did not exhibit any changes before or after damage (**Figure 5.8c**). A small increase of acetylation levels was detected compared to the LCLs: 25% and 30% for LCLs and U2-OS, respectively. The slight difference could be accounted for with the extraction methods (acid extraction in U2-OS cells) and the cell line expressing slightly different levels of acetylation. However, the results are entirely consistent.

When developing *Drosophila* larvae, H3K14ac (along with H3K23) had the highest levels of an acetylated residue (Henry et al. 2016). Both H3K14 and H3K23 were observed in having the highest proportion of acetylation in the genome for the LCLs before and after

damage. Henry et al. irradiated larvae would result in a slight variance of H3K14 acetylation levels of 1% (Henry et al. 2016). When either one of WRN or BLM were mutated, which are involved homologous recombination and NHEJ respectively, there was a significant increase in H3K14 acetylation compared to the wild-type upon irradiation (Henry et al. 2016). Even though H3K14ac level does not change in response to DNA damage in different cell lines and organisms, it still shows that it is crucial for the DDR. Previously, it was shown that H3K14ac was required for the ubiquitination of γ H2AX (Ikura et al. 2007) and (Zhu & Wani 2010). It is also important to consider that just because DSBs heterochromatin tends to be repaired via HR in an ATM-Artemis-KAP1 dependent manner, the epigenetic environment does not change to ensure chromatin fidelity, then it is possible that H3K14ac is required to remain constant before and after DNA damage to keep fidelity of more active sites of the genome and as a constantly present site for DDR with bromodomains to dock to.

Figure 5.8. H3K14 acetylation does not change post-DNA damage at either local or global levels

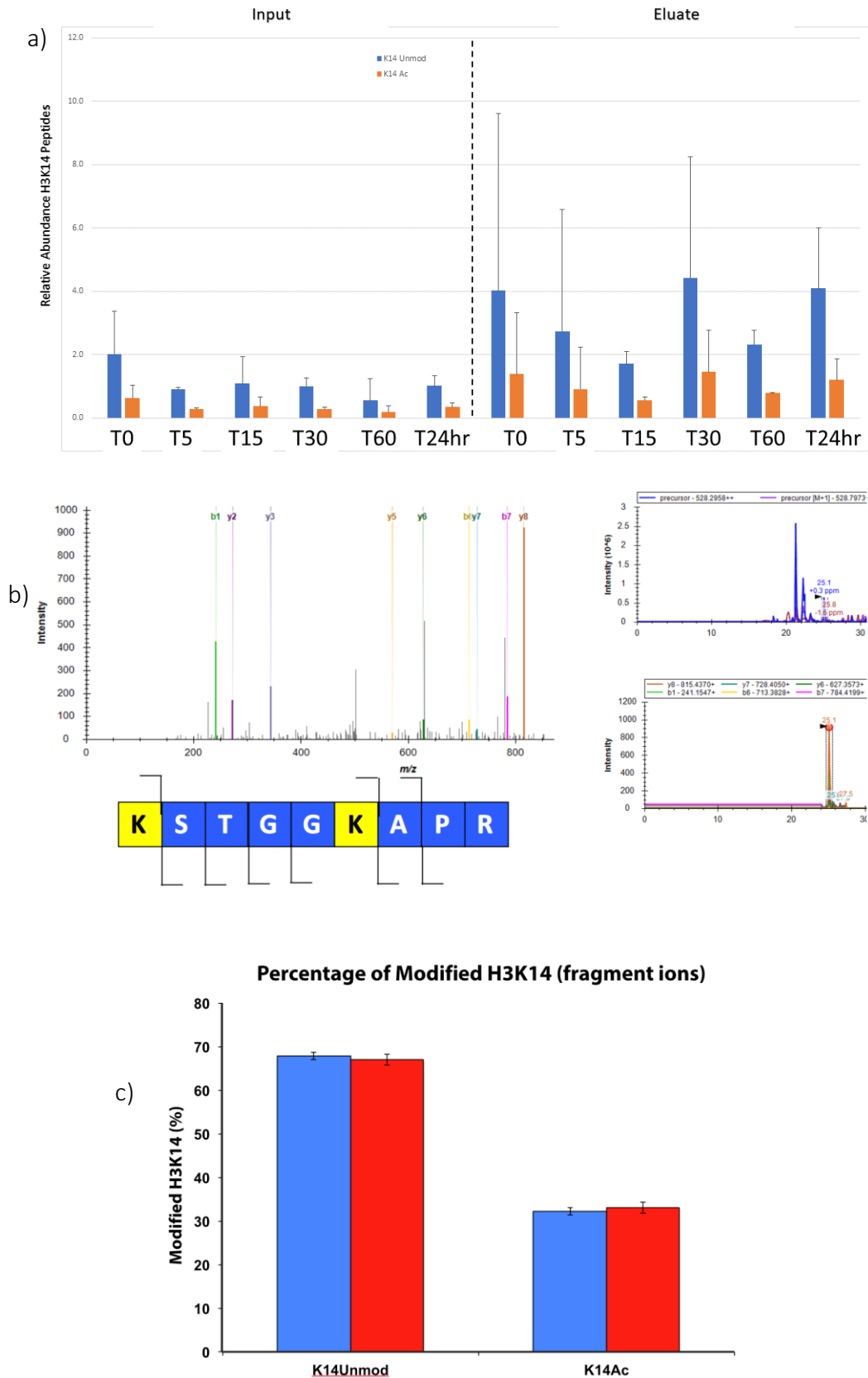


Figure 5.8. Legend H3K14 acetylation does not change post-DNA damage at either local or global levels

- a) Pulse chase analysis of H3K14 post-translational modification changes. LCLs irradiated with 20 Gy ionizing radiation. H3K14 modification percentages were quantified by measuring fragment ions for doubly charged precursors. List of peptides used for quantification:
- KSTGG[ac]KAPR
 - [me1]KSTGG[ac]KAPR,
 - [me2]KSTGG[ac]KAPR
 - [me3]KSTGG[ac]KAPR,
 - [ac]KSTGG[ac]KAPR,
 - KSTGGKAPR,
 - [me1]KSTGGKAPR,
 - [me2]KSTGGKAPR
 - [me3]KSTGGKAPR
 - [ac]KSTGGKAPR.
 - Error bars correspond to standard deviation. N = 3 biological replicates.
- b) (LHS) MS/MS fragmentation of doubly charged precursors for residues corresponding to modified KSTGGKAPR peptide. Fragment ions involved in quantitation are highlighted on the peptide sequence: b1, b6, b7, y2, y3, y5, y6, y7, and y8. (RHS) Intensity of weakest sample analysed with pseudo-SRM MS/MS
- c) Analysis of H3K14 post-translational modification changes in U2-OS before damage and 60 minutes after damage. Histones were enriched with acid extraction. U2-OS were irradiated with 20 Gy ionizing radiation. H3K14 modification percentages were quantified by measuring fragment ions for doubly charged precursors. List of peptides used for quantification identical to (a).

5.2.9. Acetylation of H3K18 increases at a local level following DNA damage

Sirt7 is responsible for deacetylating H3K18ac (Vazquez et al. 2016) (Tong et al. 2016). It has been documented that Sirt7 plays an important role in the DDR. Vazquez et al. (2016) demonstrated that Sirt7-knock out mice went through accelerated ageing. They also saw that Sirt7 was crucial for genome integrity and controlling cell cycle progression. There were increased polyploidy cells in later passages of Sirt7^{-/-} cells.

By deleting Sirt7, Vazquez et al. also demonstrate a significant decrease in 53BP1 localisation to DSBs after with induction of DNA damage. They also report that H3K18ac must be deacetylated for 53BP1 recruitment to the site of damage. To mimic the deacetylated and acetylated lysine residues, they mutated K18 to arginine and glutamine, respectively. The K18Q mutant exhibited similar levels of 53BP1 recruitment to DSBs to that of the Sirt7 knockout. On the other hand, the K18R mutant had increased levels of 53BP1 recruitment to the break site. The presence of H3K18ac may impede the interaction of 53BP1 with its docking sites: H4K20me2 and H2AK15ub. There is disruption of 53BP1 binding to H4K20me2 if the same histone H4 also contains H4K16ac. A necessary salt bridge interaction with 53BP1 on Glu 1551 has been reported. The acetylation of H4K20me2 hinders the interaction. Glu 1551 happens to be located in the Tudor domain of 53BP1 (Tang et al. 2013). There was also a reported increase in H3K18ac in yeast as upon induction of a DSB with an endonuclease (Tamburini & Tyler 2005).

If H3K18ac is increased at later time points (**Figure 5.9**), and H3K18ac has been shown to disrupt 53BP1 binding, it suggests that later time points may be going through HR-mediated repair as before. Dicer has been shown to regulate H4K16ac in the genome. Dicer which is a crucial element of heterochromatin formation in both yeast and mammalian systems (Fukagawa et al. 2004), was also implicated in the DNA damage

response. Dicer's RNase function was crucial for the foci formation of several DDR proteins. Dicer generates ncRNA in heterochromatin formation for silencing regions in a constitutive manner (Fukagawa et al. 2004). Dicer and DROSHA could generate non-coding RNA from newly transcribed RNA generated from the site of damage (Francia et al. 2012). Dicer also was shown to have an electrostatic affinity to Sirt7 (P.-Y. Zhang et al. 2016).

When HEK293T and HCT116 cells were treated with cisplatin, doxorubicin, or IR, there was increased expression of Dicer. The increased expression of Dicer resulted in an increase in Sirt7 in the cytoplasm, i.e. away from chromatin. When either one of Dicer or Sirt7 were targeted with co-IPs, they would both enrich together in the eluate (P.-Y. Zhang et al. 2016). If Dicer has a role in both an early role in recruiting 53BP1 but then also possibly suppressing its recruitment by the way on inhibiting the Sirt7 acetyltransferase activity of H3K18ac, it then raises the question of when exactly within the DDR does the switch in role take place. When ATM phosphorylates KAP1, KAP1 can repel the NuRD nucleosome and deacetylating complex from chromatin (Murray et al. 2012). 53BP1 is recruited to DSBs regardless of whether the repair is to take place via NHEJ or HR. The first response of DDR is to ligate the strands together via NHEJ, but if this takes too long, the DDR switches over to HR (Panier & Boulton, 2014). So whether the increase in H3K18ac is from late repairing breaks which require HR is difficult to say with certainty. The uncertainty is due to lysine acetylation denoting a more open environment and the risk being that with HR-mediated repair, the resection would result in nucleosomes and when they are restored they would require heterochromatin HPTM marks to be restored otherwise there could be whole chromosomal structural changes, particularly in the centromere and telomeres.

Figure 5.9. Acetylation of H3K18 increases at a local level following DNA damage

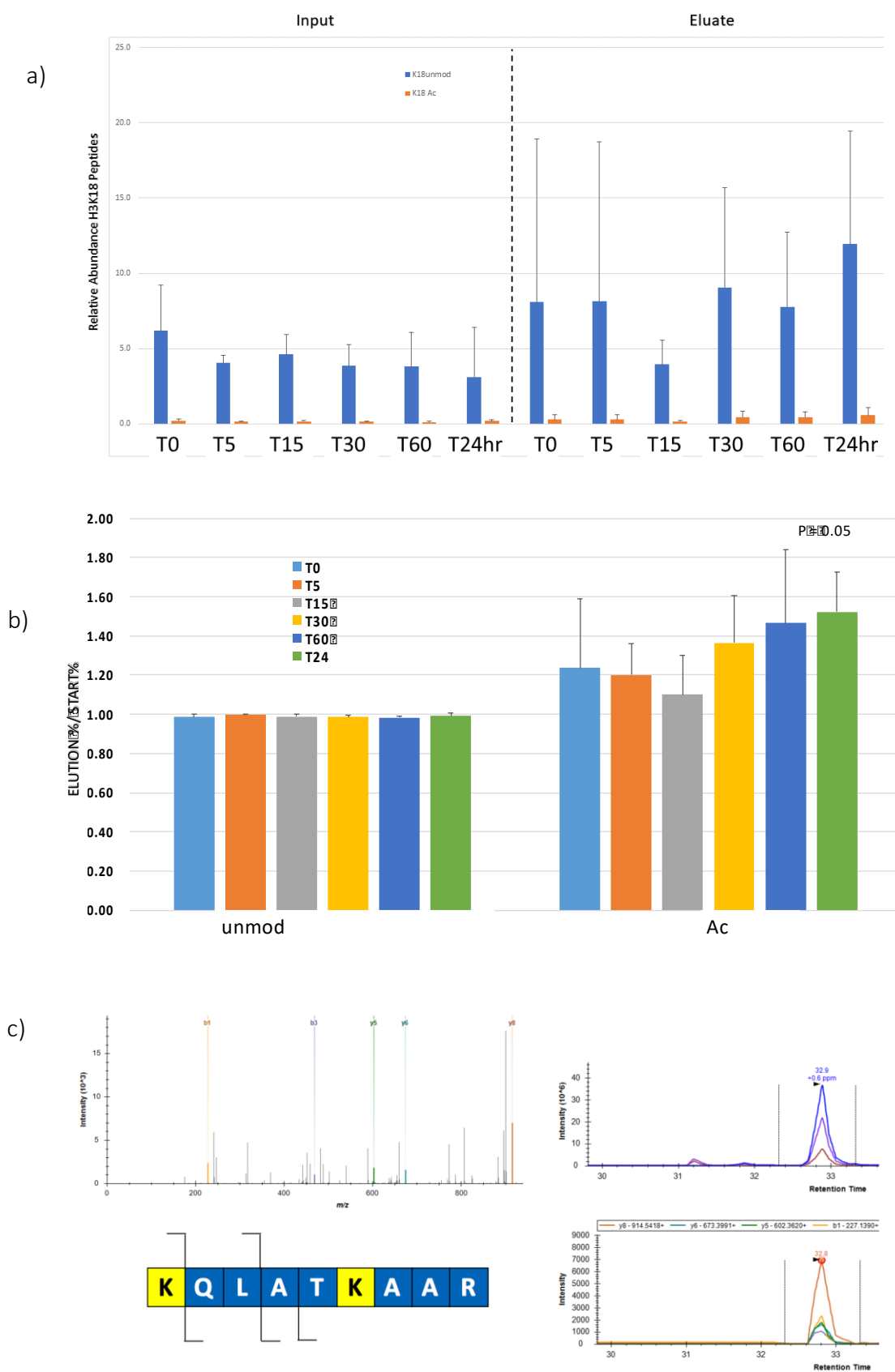


Figure 5.9. Legend Acetylation of H3K18 increases at a local level following DNA damage

- a) Pulse chase analysis of H3K18 post-translational modification changes. LCLs irradiated with 20 Gy ionizing radiation. H3K18 modification percentages were quantified by measuring fragment ions for doubly charged precursors. List of peptides used for quantification:
 - a. [ac]KQLATKAAR
 - b. KQLAT[ac]KAAR
 - c. KQLATKAAR
 - d. [ac]KQLAT[ac]KAAR
 - e. Error bars correspond to standard deviation. N = 3 biological replicates.
- b) Increased H3K18ac at the site of damage: All time points up to 60 minute indicate there is a trend but no significant difference. The 24 hour time points indicate that there is a significant difference in regards to H3K18ac at the site of damage relative to global average before damage. Two-tailed T-tests were carried out on all time points relative to their reciprocal time point
- c) (LHS) MS/MS fragmentation of doubly charged precursors for residues corresponding to modified KQLATKAAR peptide. Fragment ions involved in quantitation are highlighted on the peptide sequence: b1, b3, y5, y6, and y8.

(RHS) Intensity of weakest sample analysed with pseudo-SRM MS/MS

5.2.10. Acetylation of H3K23 increases at a local level following DNA damage

We observe an increase in H3K23 acetylation (**Figure 5.10**) similar to that of Tamburini (2005) who saw an increase in acetylation levels near damage with an HO-endonuclease site. As mentioned, developing *Drosophila* larvae have extraordinary levels of H3K23ac: highest levels of acetylation of any residue within the genome. Treatment of the larvae γ -irradiation would result in a slight variance of H3K23 acetylation levels of 1% (Henry et al. 2016). In WRN mutants, an exonuclease and helicase critical for homologous recombination, there was a significant decrease in H3K23 acetylation compared to the wild-type upon irradiation. BLM is a helicase responsible for hindering unintended homologous recombination. Henry et al. reported when BLM is mutated, H3K23 acetylation was shown to increase approximately by 1.25% (Henry et al. 2016). It has been demonstrated that H3K23 acetylation is required for ubiquitination at lysine 5 of γ H2AX, in a Tip60 dependent manner. The ubiquitination is required to expel γ H2AX at the damage site so that the chromatin can be remodelled to be accessible for DNA repair proteins.

KAT6B is a member of the MYST family of HATs responsible for H3K23 acetylation. Simo-Riudalbas et al (2015) demonstrated that when KAT6B was knocked down with shRNA, and the cells were treated with irinotecan (DNA damaging agent) they find that there is a drop off of recruitment of ATM at the site of damage and in turn a direct loss of γ H2AX. Tip60 is another member of MYST-family of HAT that has been directly responsible for activating ATM at DSBs (Y. Sun et al. 2005). It has yet to be determined whether KAT6B may be involved in the acetylation of ATM at later time points. H3K23ac was always found to be present at regions of the genome that also contained H3K4me2 and H3K36me3 in

both *Arabidopsis* and rice (Lu et al. 2015). As the damage persists and the cells' DDR fail to repair, after some time the cell will undergo apoptosis, so it could also be possible that there is rampant overexpression of apoptotic genes, especially when the cells have been treated with 20 Gy of ionising radiation. TRIM24 is a reader of H3K23ac along with unmethylated H3K4 on the same histone tail. TRIM24 co-activates oestrogen receptor α (ER- α) (Tsai et al. 2010) in breast cancer cells. ER- α is in turn known to recruit other HATs such as GCN5, CBP/p300. Also, ER- α has also been shown to actively increase transcription of tumour suppressor p53 in DNA damaged environment induced by camptothecin (Berger et al. 2012). However, drawing that conclusion leads to even more questions as ER α to recruiting: DDR proteins involved in HR, NHEJ, lesion repair mechanisms, cell cycle proteins, and apoptosis itself.

It is not possible for acetylation and methylation to occur on the same residue. H3K23 methylation in *Tetrahymena thermophila* was shown to be present in pericentromeric regions of the chromosome associated, i.e. regions related to constitutive heterochromatin (Papazyan et al. 2014). In meiosis, H3K23me3 was shown to be crucial for preventing meiotic induced DSBs by limiting access to Spo11. When EZL3, the KMT responsible for H3K23 methylation, was deleted DSBs were observed occurring in pericentromeric regions that would not normally occur. Evolutionary, it would not be beneficial to induce meiotic DSBs at regions of the chromosome crucial for structural integrity as it could result in chromosomal translocations and aberrations. However, it could be speculated that a rise in H3K23ac would result in increased HR-induced DSB, it helps explain why there is a global increase of H3K23ac across the genome: late repairing breaks require HR for repair.

Figure 5.10. Acetylation of H3K23 increases at a local level following DNA damage

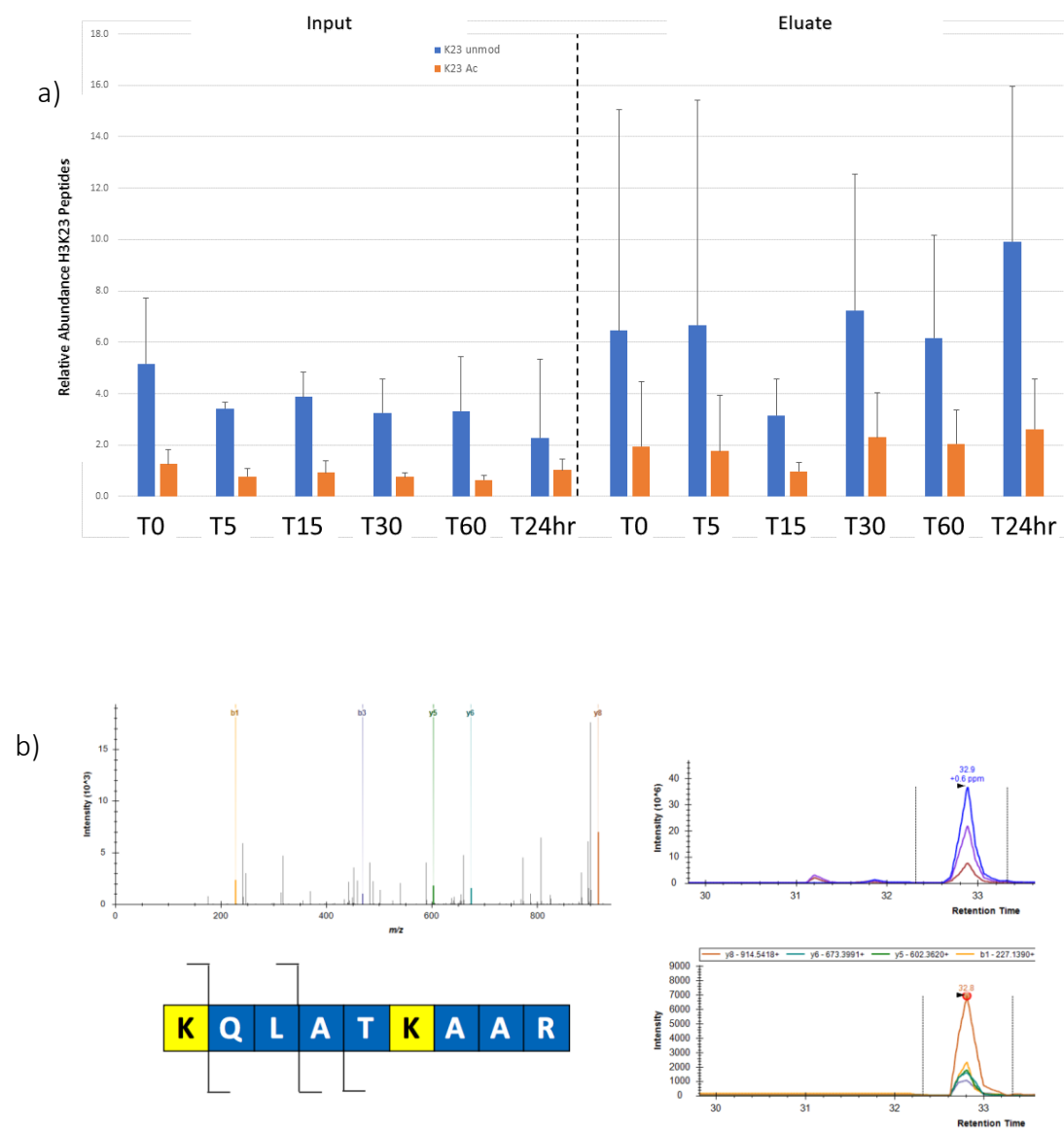


Figure 5.10. Legend: Acetylation of H3K23 increases at a local level following DNA damage

- a) Pulse chase analysis of H3K23 post-translational modification changes. LCLs irradiated with 20 Gy ionizing radiation. H3K23 modification percentages were quantified by measuring fragment ions for doubly charged precursors. List of peptides used for quantification:
- [ac]KQLATKAAR
 - KQLAT[ac]KAAR
 - KQLATKAAR
 - [ac]KQLAT[ac]KAAR
 - Error bars correspond to standard deviation. N = 3 biological replicates.
- b) Are (LHS) MS/MS fragmentation of doubly charged precursors for residues corresponding to modified KQLATKAAR peptide. Fragment ions involved in quantitation are highlighted on the peptide sequence: b1, b3, y5, y6, and y8.
- (RHS) Intensity of weakest sample analysed with pseudo-SRM MS/MS

5.2.11. Decrease in H2AZ following ionizing radiation

Histone H2AZ is a variant of H2A that differs considerably to canonical H2A, sharing only 60% homology with canonical H2A (Y. Xu et al. 2012). H2AZ has been implicated in being essential for expression of genes and is found at the nucleosomes that precede and follow the transcriptional start site (TSS) (Bargaje et al. 2012). Genes that did not possess H2AZ at their transcriptional start site resulted in decreased expression compared to genes that had a higher proportion of H2AZ near their TSS, i.e., there is a correlation of H2AZ proximity and gene expression (Bargaje et al. 2012). H2AZ occupancy at transcriptional start sites was also observed in *Arabidopsis* (Coleman-Derr & Zilberman 2012). H2AZ has been reported to be recruited to sites of DNA damage and quickly overturned within five minutes. The irradiated LCLs exhibited a drop in H2AZ levels at the site of damage (**Figure 5.11**). A higher resolution time course would be required; a more intricate set up that omits centrifugation as by the time the cells are spun down the time-sensitive events have already taken place before the cells could be snap-frozen. Xu et al first reported an exchange of H2AZ at the site of damage in p400 ATPase-dependent manner (Y. Xu et al. 2012). They induced a DSB at a single point in the genome with p84-zing finger nuclease, and their ChIP data showed that there was an exchange of H2AZ very quickly at the damage site. There was no loss of H2AX, so H2AZ was not swapped for γ H2AX at the site of damage. In the presence of damage, histone H2AZ would spread approximately 10-50KB away from the DSB in either direction. H2AZ exchange is critical for HR-mediated repair (Y. Xu et al. 2012). Anp32 (H2AZ unique chaperone) was also showed to be essential for evicting H2AZ from the genome (Gursoy-Yuzugullu et al. 2015) (Alatwi & Downs 2015). When Anp32 was knocked down, or its H2AZ chaperone activity was inactivated, there was reduced acetylation of histone H4. The loss could be directly

attributed to the presence of H2AZ at the DSB which was reported to be transiently present. Depletion of Anp32 also resulted in a decrease of HR, as H2AZ could not be evicted. Rad51 and RPA foci were reduced by more than half when compared to wild-type. Cells that had both H2AZ with Anp32 co-depleted resulted in rescuing of HR-mediated repair. Consequently, Rad51 and RPA foci formation were rescued to wild-type levels. H2AZ increases and eviction seem to occur as an early response to DNA damage. What we see from our data is that there is a loss of H2AZ from the site of damage by about fifteen minutes. The issue is that H2AZ levels appear to increase again. As mentioned, before γ H2AX formation takes place in more transcriptionally active regions of the genome (Cowell et al. 2007). Thus the transient increase in H2AZ levels may be explained by global levels of H2AZ not increasing but nascent H2AZ along TSS being pulled down along with the γ H2AX, thus inflating the value. At the earlier time points, we are actively seeing the H2AZ being evicted. The presence of H2AZ at the site of damage is very short lived which is a problem in this assay with DOE: H2AZ is fully removed by ten minutes after the induction of damage (Gursoy-Yuzugullu et al. 2015) (Alatwi & Downs 2015). Only one-time point was collected either side of the ten-minute mark. To investigate levels of H2AZ at the quickest of time points, we would need to collect short time points and thus not allowing the cells to recover as much.

Figure 5.11. Decrease in Histone H2AZ following ionizing radiation

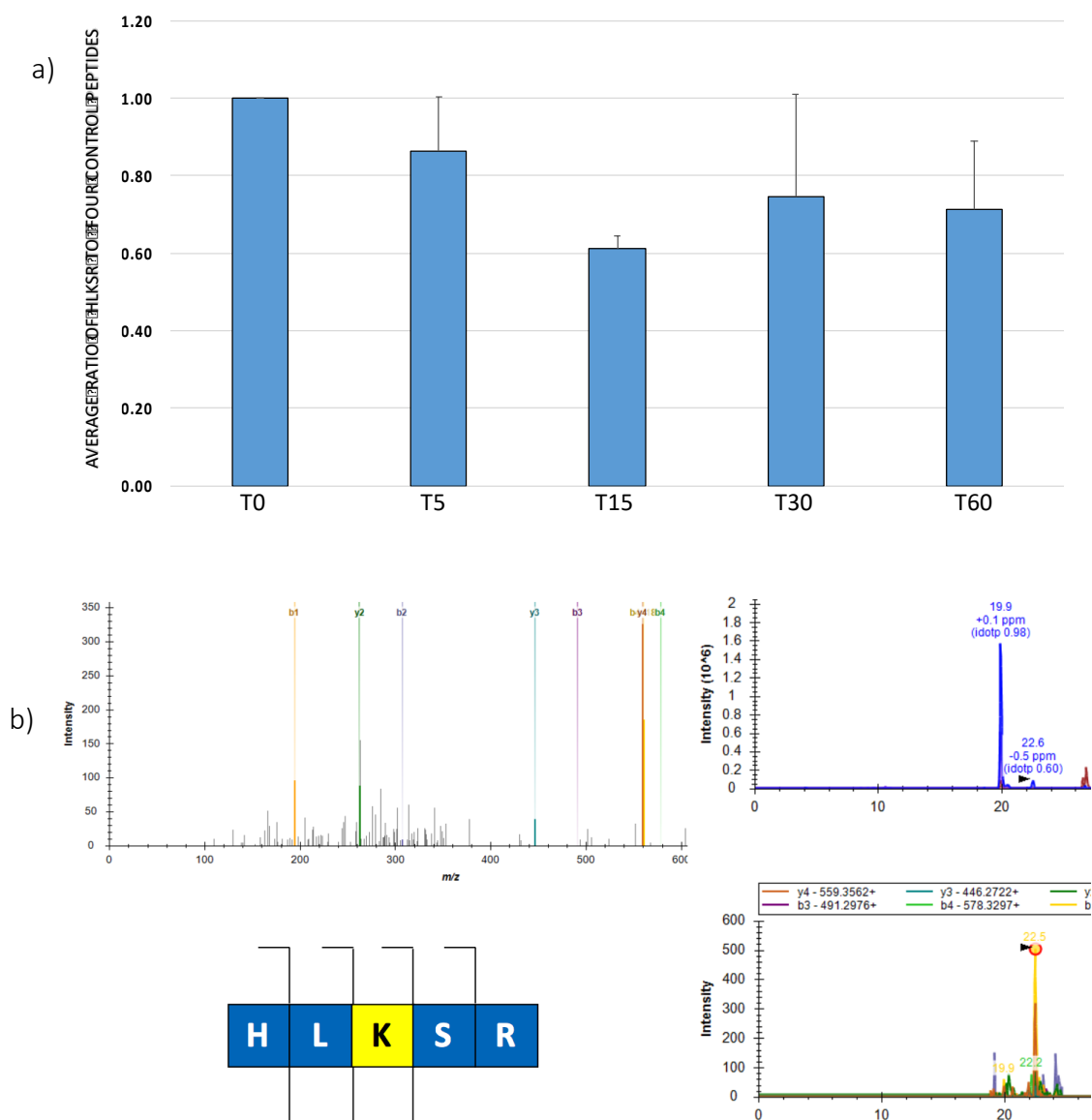


Figure 5.11. Legend Decrease in Histone H2AZ following ionizing radiation

- a) Pulse chase analysis of H2AZ abundance the site of damage. LCLs irradiated with 20 Gy ionizing radiation. H2AZ modification percentages were quantified by measuring fragment ions for singly charged precursors and measuring H2AZ against control peptides. List of peptides used for quantification:

- a. YRPGTVALR [Histone H3 41-49]
- b. YQKSTELLIR [Histone H3 54-63]
- c. VTIMPKDIQLAR [Histone H3 118-129]
- d. DAVTYTEHAKR [Histone H4 68-78]

Error bars correspond to standard deviation. N = 3 biological replicates. Transient decrease at of H2AZ at the site of damage: At the 15 min time point relative to the 5 min there is a decrease indicate there is a trend but no significant difference.

- b) (LHS) MS/MS fragmentation of singly charged precursors for residues corresponding to modified HLKSR peptide. Fragment ions involved in quantitation are highlighted on the peptide sequence: b1, b2, b3, b4, y2, y3, and y4. (RHS) Intensity of weakest sample analysed with pseudo-SRM MS/MS

5.2.12. Undetected Known DDR-related Histone post-translational modifications

In our studies, there were certain histone post-translational modifications that have been reported to change in response to damage that we failed to detect, let alone quantify. These include acetylation of H3K56, H3K4 methylation, phosphorylation of H3S10, and ubiquitination of both H2AK13/K15 and H2AK118/K119.

In previous studies, H3K56 acetylation has been shown to be reduced upon induction of DNA damage in various species: *S. cerevisiae*, *P. falciparum* and mammalian systems (Yuan et al. 2014) (Gupta 2016) and (Tjeertes 2009). In yeast, changes in H3K56ac levels were detectable via mass spectrometry at both the site of damage and genomic levels (Ozdemir et al. 2005). In addition, Unnikrishnan et al. developed a yeast strain that was transformed with a non-integrative plasmid that was designed to be pulled down for mass spectrometric analysis of the histone proteins associated with the plasmid (Unnikrishnan et al. 2010). However, attention is focussed on to mammalian cells, reported changes in H3K56ac levels as a result of DNA damage have always been with immuno-based methods (Tjeertes 2009) and (Yuan et al. 2014). When Zhang et al. carried out a post-translational modification proteomic screen on histones prepared from calf thymus cells, they received data that was too ambiguous to assign actual acetylation to the H3K56ac (L. Zhang et al. 2003). In a screen to determine how universal is the histone code across species, Garcia et al. (2006) demonstrate the discrepancy in the mass spectrometric detection of H3K56ac across eukaryotes directly. In both *S. cerevisiae* and *T. thermophila*, H3K56ac was detected in proteomic and immuno-based assays. However, even though there was a detection of H3K56ac in human cells with both proteomic and immuno-based assays, it was detected with low abundance. In mice cells, they do not

even detect H3K56ac in either proteomic or immuno-based assay. It could be possible then, that even though H3K56ac has been a highly characterised HPTM, especially in regards to its role in transcription and DNA damage, it is possible that it is only present in the genome at very low levels. As we saw in chapter 3 with regards to H2AXY142 phosphorylation, it is possible that some modifications are present at levels below the detection threshold of the instrument. Even though immuno-based assays require a degree of caution, they do have the advantage of not requiring copious amounts of the epitope compared to mass spectrometry which can be sensitive to sample preparation techniques or there not being enough of the peptide of interest sample in the first place.

In asynchronous cells, it is quite difficult to determine whether there are changes in H3S10p since the modification occurs mainly in mitosis and within an asynchronous population, only a minuscule proportion of the cells would be present in M-phase (A. K. Sharma et al. 2015)(Tjeertes 2009). To overcome the cell cycle artefact masking potential changes in H3S10 phosphorylation, Sharma (2015) and Tjeertes (2009) resorted to synchronising the cells to G1 cells and inducing damage. Sharma et al. synchronised with serum starvation and induced damage with either 2.5Gy and 15Gy, and with both treatments, they see an inverse correlation of H3S10P with γ H2AX, where there is a rapid decrease in H3S10P up to 8 hours upon treatment with DNA damage. Then levels of H3S10P start to recover while levels of γ H2AX started to drop. At 24 hours there is full recovery of H3S10P. However, even though Tjeertes et al. (2009) saw an increase in H3S10P during mitosis, they were not able to detect changes in H3S10P levels when synchronised cells were treated with phleomycin. Also, there were no changes in phosphorylation at H3T11, H3S28, and H3.3S31. We were not able to detect H3S10P

levels, and this could be attributed to a few points that concur with Tjeertes and Sharma: this study used cells that were asynchronous, so levels of H3S10 phosphorylation would have been very small, to begin with. H3S10P and γ H2AX have an intrinsic nature to not occupy the same nucleosomes or genomic regions thus even if there were any changes to phosphorylation of H3S10P following damage, nucleosome enrichment via γ H2AX would not pick up the changes compared to a different histone such as H4. Even if we do not take into consideration that Tjeertes saw no change of H3S10P levels following phleomycin treatment, there is still the issue of ionisation efficiency. As previously shown, adding a phosphorylation modification to a peptide can significantly reduce the ability for the MS/MS to even pick up the peptide (**Figure 3.1**).

The H3K4 data was problematic for various reasons. One particular issue with the H3K4 peptide was that it was incredibly hydrophilic and would be the first to be eluted from the nano-LC into the Orbitrap. The early elution becomes a hindrance to quality data collection as the retention times for the different masses would drift slightly. Having many overlapping retention times or peptides with low abundance would be lost in the noise. We had low detection of H3K4 peptides. The low detection of H3K4 modified peptides could be attributed to H3K4 peptides needing to be derivatized to prevent trypsin cutting at unmodified or monomethylated H3K4 lysine residues. When propionic based derivatization is added to the sample preparation before proteomic analysis, another step is added where peptides can be lost. Maile et al. reported similar issues with detecting with H3K4 PTM (Maile et al. 2015). Maile et al. demonstrated that sample preparation preceding proteomic analysis leads to unequal recovery of identical peptides where the single difference are different PTMs on the same residue i.e. a different

residue on the same histone peptide will result in large discrepancies in recovery and ionisation efficiencies (**Figure 3.1**). They generated synthetic peptides corresponding to the first 18 residues of histone H3, except for the K4 residue which either possessed no modification, mono-, di-, trimethylation, acetylation, with and without phosphorylation at threonine 6. They noticed that the synthetic peptides that carried one of the H3K4me2, H3K4me3 with either/or phosphorylated H3T6 would only be recovered at 10% levels of equivalent peptides containing no modification or H3K4ac. H3K4me1 would have a recovery at twice the levels of both unmodified and acetylated H3K4. Low ionisation efficiency and recovery after propionic derivatisation does not help with the proteomic assay of modifications that are present at very low levels within the genome. H3K4me3 has been characterised to be present at the start of transcriptional start sites (Ji et al. 2015). However, these sites are very few when taken into consideration the genome as a whole. H3K4 methylation has been implicated in the DNA damage response, so it is unfortunate we were not able to apply the γ H2AX-IP-MS at this round. Maile et al. (2015) demonstrated that switching from a propionic based method to a PIC based method in the second step of derivatisation (**Methods and Materials**) would allow for an increased relative abundance of the more hydrophobic peptides, resulting in improved ionisation efficiency. However, if this approach was to be utilised for HPTM extracted from an in-vivo sample, the two derivatisation methods might have to be used concurrently to allow for weighting of ionisation efficiency discrepancies that result from using two different reagents on the same peptides.

In our study, we do not detect any ubiquitination with a degree of certainty. When trying to attribute the peaks to masses corresponding to known ubiquitination sites in the mass

spec-quantifications analytical software, the signal-noise ratio was far too low. At least with the modifications, we quantified above, there were clear peaks that were way above background levels to pick from. Ubiquitination of H2AK13/K15 has been shown to be essential for 53BP1 loading at the site of damage. 53BP1 recognises H2AK15 ubiquitination via its UDF motif. Following MDC1 recruitment to the site of damage in a γ H2AX dependent manner, the E3 ligases RNF8 and RNF168 are subsequently recruited to the DSB. As a result (Panier & Boulton, 2014). RNF8 and RNF168 are responsible for ubiquitinating H2AK15 (Panier & Boulton, 2014). The H2AK118/119 residues are ubiquitinated in a manner dependent on various members of Polycomb group complexes (PcG), particularly those from PRC1 and PRC2. RNF8 works with UBC13 to monoubiquitinate H2AX and H2A (Gieni et al. 2014). UBC1 will When Ring2 and EZH2 activity was suppressed by either knockdown or loss of function there was increased sensitivity to ionising radiation (Gieni et al. 2014). Ring2 and EZH2 are part of PRC1 and PRC2, respectively. Ubiquitination of H2AK119 can occur at both sites of damage and at chromatin that has been repressed in a PRC-1 silenced manner. PRC1 silences the *Drosophila* transcription factor Ubx indirectly by via the ubiquitination of H2AK119 (H. Wang et al. 2004). The ubiquitinating activity of RNF168 is required for the accumulation of K63 polyubiquitin at H2AK15. The K63 polyubiquitin site will, in turn, be recognised by RAP80, BRCA1 recognises RAP80 at the site of damage and initiates HR-mediated pathway (Gieni et al. 2014).

So if ubiquitination of lysine residues has been explicitly shown to be involved in the DDR, then why do we not detect any ubiquitin with an element of certainty that allows for quantification. Again, this is not due to one particular reason but a combination of

multiple factors. In the nuclear extraction step of the sample preparation, NEM is added to ensure there is no extracellular deubiquitination of ubiquitinated sites. However, during method development, it was observed that Mnase would have its nuclease activity impeded in the presence of high concentration. So to resolve this issue, we transiently decrease the levels of NEM and other inhibitors so that the Mnase can digest the DNA in the chromatin. Another factor is that with Mnase digestion the temperature must be approximately 37 °C. However, NEM is highly sensitive to increased temperatures and will completely degrade within about half an hour at room temperature. Therefore, during Mnase digestion step, NEM should be replaced regularly while the digestion is taking place to prevent loss of nascent ubiquitination. The next issue could be found in the sample preparation following the immunoprecipitation. For a bottom-up approach in proteomics, we utilise trypsin to digest the protein-enriched samples so that peptides fragment better and masses can be more accurately obtained with the Orbitrap. However, with the pseudo-SRM targeted method that were utilised, the expected mass needs to be known in kD up to four significant figures, prior to injection. The mass of the peptide can be determined from the sequences that can be obtained from many online resources such as Uniprot. The C-terminal of ubiquitin ends with the amino acid sequence: KESTLHLVLRIRGG (G. Xu & Jaffrey 2013). The glycines will be covalently bound to lysine. When trypsin cuts to completion, only the last two glycines will remain as trypsin digests at arginine and lysines that do not possess large modifications. However, it is possible that incomplete trypsin digestion can lead a sequence of LRGG remaining on a ubiquitinated lysine.

Regardless, the presence or absence of LRGG- sequence along with GG- sequence does not complicate statistical analysis, as their signals could be pooled together. The levels of H2A ubiquitination could be artificially deflated if all the different possible permutations of ubiquitination tryptic-artefacts are not accounted for during a targeted proteomic screen. Of course, when one is carrying out a trypsin digestion they do not do so with the intention of expecting there to be partial trypsin digestion.

The final factor that can be analysed is over-propionylation resulting in unintended side reaction with the peptide at the derivatisation. Apart from analysing γ H2AX, all the other residues that were analysed were prepared with a derivatisation step. The peptides are propionylated before and after trypsin digestion. The first propionylation step is used to ensure equal lengths of an amino-acid sequence are produced as trypsin will cut at lysine residues that are not modified or only possess a monomethylation. The second is to ensure that new N-termini are also propionylated. Meert et al. (2016) demonstrated that there is a fine line between over-derivatisation which can result in nonspecific propionylation at serine, threonine and tyrosine residues, and under-derivatisation, which will lead to irregular peptide sequences for quantification. Even though there was a step that was used to ensure reversal of over-propionylation with the addition of hydroxylamine (HA). The HA and propionylation can then interfere with the trypsin digestion, so there is a fine balance of not overshooting with either.

An alternate derivatisation described by Fiedler et al. (2013) was to derivatise the ubiquitinated lysines directly by swapping out the ubiquitinated lysines with a glycosylated lysine. The glycosylated lysine reaction could be used for identification of new ubiquitination site along with quantification of the ubiquitination site. The new lysine residue has very high chemical mimicry of the old residue after successful trypsinization.

This technique could be utilised before trypsinisation without the requirement of a propionic derivatisation if you are only interested in ubiquitination.

Chapter 6: Concluding Discussion

The results of this thesis build on of proteomics by developing methods suitable for answering questions about changes in HPTMs in a damaged chromatin environment, which was not previously attempted at such scale. A brief synopsis of the results from the preceding chapters will be outlined, which will be followed by a holistic discussion into the significance of the method developed, outlining where it could be improved and what possibilities lie in the future with the method.

6.1 Summary of Results

The initial aim was to carry out a mass spectrometry-based global analysis of HPTM changes that occur as a result of the DDR. Established acid extraction protocols were used to prepare histones. However, it became abundantly clear that, apart from detecting changes in γ H2AX levels, acid extraction followed by global quantification was not very useful for detecting nuanced changes in other HPTMs which are DNA damage-dependent (**Chapter 3**). It was theorised that since most HPTMs do not going to give a strong damage response signal at a local level, it would be even more difficult to detect at a global level as many factors could affect the detection of changes in HPTM levels. For example, H3K9me3 can be found in the centromere as part of the constitutive heterochromatin (that is localised to that region of the chromosome). If H3K9me3 levels were found to either increase or decrease by 1% after damage, such a subtle change would likely be masked by centromeric H3K9me3. On the other hand, detecting changes in γ H2AX levels was more straightforward. Levels of γ H2AX before and after damage were quantified using the same method. Subsequently followed up the analysis by quantifying the

abundance of H2AX compared to other H2A variants in multiple cell lines. Even though there was a slight discrepancy amongst the cell lines, the differences were still within a range that could be attributed to cell line specific abundances. However, when mass spectrometry quantification of phosphorylated H2AX-Y142 was attempted, it transpired that earlier studies had possibly exaggerated the levels of phosphorylated H2AX-Y142 due to the antibodies used against the modification (Xiao et al. 2008). The basic levels of Y142 phosphorylation were below the levels of detection with our methods. To ensure it was not a technical or instrumental error, we used synthetic peptides that corresponded to a trypsinised C-terminal H2AX peptide which carried the phosphorylation on the Tyrosine 142 residue. We were able to detect it even when spiked into an acid extract mixture from U2-OS cells. As a result of not detecting more changes that occur because of the DDR, we chose to develop a method that could detect nuanced changes that are more proximal to the site of damage. In **Chapter 4**, several avenues were chosen for developing such a method. At first, we attempted to use a DDR protein that localises with damaged chromatin. 53BP1 was selected but even though 53BP1 was successfully tagged with FLAG, levels of 53BP1 could not be detected that would be sufficient for a proteomics study of chromatin dynamics during the DDR. We then opted for the approach of pulling down nucleosomes directly. γ H2AX was chosen as it is a well-characterised marker of DNA damage. Since we were introducing an immunoassay step to the proteomics-based method, we tested the feasibility of this approach with several control steps to ensure the best antibody was chosen. We devised a system that would enrich for phosphorylated H2AX and therefore pull-down chromatin specifically from damaged regions. With this new system, which we called γ H2AX-IP-MS, we were then able to analyse HPTM changes at a more nuanced level. In **Chapter 5**, we detected

enrichment of γ H2AX signal but not H2AX as a whole in comparison to other variants, i.e. the ratio of H2AX enriched was similar to that of a global enrichment. We eradicated non-specific binding of histones to the γ H2AX antibody in **Chapter 4**, and the putative explanation for this result is that the long chain of nucleosomes that we used to IP from was populated with other histone variants rather just H2AX alone. This issue explains and sets the stage for most of the data that was accrued the first time around. What was important to see, even if the system did not work as expected at its maiden attempt, was that there is still a difference between local and genomic levels for HPTMs before and after damage. We detected small changes in heterochromatin dynamics in H3K9me2, no change in H3K14ac levels in response to DNA damage, and an increase in H3K18ac at the 24-hour time point. There also seems to be a small inclination for H2AZ to be evicted from the nucleosome. This difference in levels highlights that the system is feasible and showing glimpses of its potential. Once the technical issues that inevitably follow any new novel methodology have been removed from the system, γ H2AX-IP-MS can become a powerful utility to analyse HPTMs in the light of DNA damage. The system would also not just be limited to the DDR but could work in other environments such as transcription as it was designed to be modular from the onset of conception.

6.2 Design of Experiment Audit

In this section, we explore possible alterations to the method that could result in cleaner signal and increased yield of data. These improvements to the system are intended to take place both before and after the IP, as the IP itself has been optimised.

6.2.1 LCLs: Use more cells and a lower dosage of γ ray

When we started irradiating cells, we decided to treat the cells with 20 Gy, with the idea being that we would get a large response from the cells in regards to the DNA damage. In hindsight, this level of irradiation would have been too much for the cells to cope with. 20 Gy is above and beyond being excessive as 2-4.5 Gy is lethal. As a response to this, we should lower the dosage to a more physiological response which would be less than 2-4.5 Gy (Levin et al. 1992; Mole 1984). The question then arises of what would the academic merit even be to dose cells with such large amounts of radiation? Most of the world's human population is subjected to an annual dose of radiation ranging between 1-10 mGy, which comes from background radiation (**World Health Organisation 2016**). In some rare circumstances, there are scenarios where the dosage of radiation would exceed the average. For example, during radiotherapy, a tumour could be subjected to an IR course of 1.8-3 Gy over a 6-8 week period, which would give a final accumulated dose of up to 75-80 Gy (**Prise 2017**). Some residents who were close to the Chernobyl Power Plant disaster had their thyroid glands exposed to several Gy of ionizing radiation and some of the powerplant workers received an excess of 240 Gy over a two year period. As a result, they died shortly after as a consequence of acute radiation sickness. (**World Health Organisation 2016**). Apart from these special rare cases, it would be highly unusual for a cell to be subjected to such high doses of radiation. Not only would such large doses of radiation result in the severe fragmentation of the DNA strands, it would also lead to many of the nuclear proteins being impacted by the radiation. Setting aside that damaged proteins could build up into hazardous material such as amyloidosis (**Pepys 2006**) the other issue is that the repair proteins themselves would be compromised. As mentioned, radiation can cause damage to the proteasome in addition to the DNA strand

(Radman 2016). Even if there was no damage to DNA itself, if the proteasome was wiped out, the cell would not be able to survive because even though the instructions remain intact, there are no readers available to read the instruction (Radman 2016). It then stands to reason to question the efficacy of radiating the cells with large doses.

The other factor to take into consideration when using large amounts of radiation is that it will result in a staggered response from the different actors in the DDR. It has been shown that the rate of phosphorylation of H2AX will occur at a much faster rate compared to other factors coming into play such as 53BP1 (Asaithamby & D. J. Chen 2009). It is important to note that the DDR proteins are not drawn from an infinite pool of proteins anticipating to be recruited. There is a limited amount of DDR proteins, so blasting the cells with an immense amount of radiation will not give an accurate indication of what is taking place at the site of damage if the DDR proteins have not been given a chance to respond to damage in a canonical physiological manner. For example, if at 20 Gy the cell has ten times as much γ H2AX compared to a cell that has been irradiated with 2 Gy but the amount of 53BP1 and chromatin modifiers were saturated at 2 Gy, then all that would have achieved is the generation of more noise in the pull-down. Therefore, the more appropriate solution would be to lower the dosage of radiation by tenfold for example and increase the number of cells ten times. The issue with this readjustment of dosage to cells would be handling a significant amount of cells. However, it would mitigate against generating new γ H2AX that are at sites that are lagging in the other aspects of the DDR.

6.2.2. Exclusively Use S1; replenish S2 with more Mnase

S1 is the initial supernatant that is mainly composed of mononucleosomes and dinucleosomes, upon treating the nuclei with Mnase. The S2 is made up of nucleosome chains that contain several nucleosomes: 7-8. In **Chapter 3** we show that H2AX variance has no more than 5% of all H2A variants. A rough estimate would show that upon digestion, nucleosome chains that are eight nucleosomes long are generated. There would be 16 H2A histones in that chain. If there is non-symmetric deposition, then 1 in 16 histones (two histones per nucleosome) equates to 6.25% which is similar numbers to what we saw with genomic levels of H2AX. With hindsight, using chains of 8 nucleosomes meant that the proportion of histone H2AX was not dissimilar to that of global levels, which explains the small increase in H2AX levels we saw in **Chapter 5** but a markedly more substantial increase in γ H2AX levels in the eluates. This was an error in the DOE as the initial intention was to obtain as much histone signal as possible. It also seemed logical at the time that nucleosomes that may not possess an H2AX variant would still exhibit HPTM changes as a result of the DDR. In future experiments, the best approach would be to only use the S1 supernatant rather than pooling the S1 and S2 together.

6.2.3. Derivatisation

A potential issue is that the derivatisation step was working on different amounts of protein for both the eluate and input. For the in-gel derivatisation step, 5 μ g of protein was used. However, the final eluate after a γ H2AX IP would be several magnitudes smaller than 5 μ g. This difference in amounts could result in either over-derivatisation of the eluate or under-derivatisation of the input. Downstream of this, the mass spectrometry data could lead to possessing different patterns of modified proteins as a result of the inconsistency. Both samples were derivatised in-gel, and this can also lead to asymmetric

losses of the samples. To account for these possible discrepancies, masses (m/z) could be targeted that demonstrate over or under derivatisation. The idea is to target an as high derivatisation as possible to eliminate any small inconsistencies swinging the data one way or another.

6.2.4 Controls to test region of the genome for Mnase activity

Another possible shortfall for Mnase is that it may not be digesting uniformly across the whole of the genome. Mnase cuts regions of the genome slower when those particular region are less accessible (Mieczkowski et al. 2016). Therefore, to account for the variability of the Mnase digestion of chromatin, there are several control tests that could be added to ensure the Mnase digestion's lack of uniformity does not affect the results downstream.

6.2.4.1. Acid Extraction

Acid extractions are useful for extracting all of the histones in an indiscriminate fashion. When carrying out a nuclei isolation, some nuclei would be aliquoted for an acid extraction. Acid extractions are a better control for global HPTMs as the acid extraction can capture histone from all parts of the genome unlike that of Mnase alone. If an acid extraction from equivalent time points were run with the idea to look for the same transitions, it would give a more accurate account of the global levels of HPTMs.

6.2.4.2. H2AX

In the γ H2AX-IP-MS method, only nucleosome chains that contained γ H2AX were enriched. However, if under the same conditions, some of the samples were immunoprecipitated with an α -H2AX antibody, the α -H2AX antibody would act as a control for where H2AX is deposited in the genome. This approach would give better data of the proportion of H2AX that is pulled down, rather than extrapolating the same data from γ H2AX-MS-IP. The direct data would give a more accurate reading of where H2AX is

deposited. If the γ H2AX-IP were cross-checked against the H2AX-IP, it would highlight whether there is discriminate regional phosphorylation of H2AX. For example, with a H2AX-IP-MS, if we were to see increased levels of euchromatin or heterochromatin compared to global levels, it would highlight that there is discrepancy in the regions that are being analysed globally and locally.

6.2.4.3. Carrying out H4-IP-MS

Histone H4 only has one variant (**The UniProt Consortium, 2017**) which is advantageous for this application because it would be deposited into every nucleosome, therefore, it can be found in every region of the genome. If the same procedure of immunoprecipitating γ H2AX as described previously were carried out on Histone H4, it would result in better representation in the region of the genome where Mnase has had its most efficacy. Histone H3 does not possess much sequence differences within its variance. However, using histone H4 exclusively would eliminate all doubts because with Histone H3 there is always the question of whether the antibody of choice is discriminatory towards H3.1 instead of H3.3 for example. Therefore, if Mnase is able to digest everywhere uniformly, H4-IP-MS should give the same data for HPTMs as AE-MS but practically the data will not be the same and Mnase will instead cut more frequently at open sites. This would not be a test of whether Mnase cuts at more open sites or not. That has already been documented (**Mieczkowski et al. 2016**). This would be an assay in measuring Mnase cutting variance of the sample from one batch to the next. It is also possible to use this method to normalise γ H2AX data at a global level in comparison to acid extraction.

6.3. Future Applications

6.3.1. H2AX Symmetry

Histone symmetry is when two copies of a single variant of a histone occupy the same nucleosome. Similarly, HPTM symmetry would be when both of the same variants within a single nucleosome carry the same modification. To answer the question of whether a single variant occupies a nucleosome or there are multiple variants deposited in the same nucleosome would be very difficult if not outright impossible with immunoassays. This is an example of where proteomics can outright answer questions that are not possible with immunoassays. Voight et al. (2013) used mass spectrometry to demonstrate that H3K27me3 is symmetrically modified on some nucleosomes and asymmetrically modified on other nucleosomes. They also demonstrated that antagonistic modifications could not occupy the same histone tail, but they could occupy the same nucleosome by being situated on sister histones.

To add to this, with the use of single-molecule imaging, Shema et al. (2016) found that the overwhelming majority of nucleosomes are asymmetrically modified. The authors only detected one H3K4me3 on histone H3 in the nucleosome, and they found the same result for H3K27me3. These studies indicate that HPTM asymmetry is a phenomenon with potentially interesting biological functions. Therefore, symmetry in both nucleosome deposition and modification of tails can affect the signalling of the histone code. Many questions would arise in regards to H2AX: would H2AX be deposited with a twin H2AX in the same nucleosome or be deposited alongside another H2A variant? If H2AX is not deposited alongside another H2AX, is the other H2A deposited a specific variant or is H2AX randomly deposited with any other variant? H2AX has also been shown to interact directly with 53BP1: would the number of H2AX affect the docking of a 53BP1

at a given nucleosome? If the answer is that there is symmetric deposition sometimes, but the genome exhibits asymmetric deposition at other sites, would there be a correlation between where the symmetry and asymmetry take place? For example, would more open sites of the genome be more susceptible to symmetric deposition, i.e. more dense with H2AX? For this experiment, both α -H2AX and α - γ H2AX could be used. If an α -H2AX-IP were carried out on an S1 so that only mono-dinucleosomes were enriched, if H2AX were deposited symmetrically then the percentage signal for H2AX would be between 50-100% compared to total H2AX variance. Similarly it is not difficult to imagine that if H2AX is deposited symmetrically will it also phosphorylate symmetrically? The stoichiometry could have effects on 53BP1 docking. 53BP1 forms a dimer at the site of damage. Does each molecule require its own phosphorylated H2AX and ubiquitinated H2A to dock to? It is also possible that only one of each modification occupies the same nucleosome, i.e. a nucleosome during the DDR will host one H2A K13/K15 ubiquitination and one γ H2AX.

6.3.2. Modular DOE

When this system was being conceptualised, it was intended to be modular from the onset. Just because the system was optimised with γ H2AX, it does not mean that it could not be reproducible with any characterised antibody for a particular modification, variant or protein of interest. As long as the validation as described in **Chapter 4** is repeated for each new protein of interest, the system could be applied to other questions that are not just exclusive to DDR. Examples include assessing HPTM changes with an antibody for H3S10 phosphorylation or H3K36 methylation. The system is not just limited to HPTM. It could be used to assess different histone variants such as H2AZ. In conjunction with SILAC, H2AZ could be analysed to determine where the histones migrate to after being

rapidly evicted from the nucleosome. It would also answer the question of whether new H2AZ are generated, or old ones are recycled during the DNA damage response. Assessing histone variants could be repeated for histone H3.3 or macroH2A.1.

6.3.3. Late Repairing Breaks

Goodarzi et al. (2008) showed that in ATM-mediated repair of breaks occur, take place in the more difficult to access regions of the genome, i.e. heterochromatin, in a CAP1 dependent manner. The heterochromatin specific breaks are more likely to be repaired via HR. A γ H2AX pulldown should show that there is increased H3K9me3, which is a marker of heterochromatin. Since HR is taking place, there should also be increased H3K36me3 signal. The question then arises of whether these modifications are old and were always present in the genome or whether they are new modifications which occurred sometime after the DNA break. The γ H2AX-IP-MS system could be coupled with SILAC to answer questions in regard to the origin and timescale of the modifications: it would be possible to discern whether the modifications were old or new, i.e. did the breaks occur in regions of heterochromatin or did the region become heterochromatic ad-hoc. It would also help to clarify whether the histones were present prior to the break or if the histones were deposited after damage taking place. These results would then lead to mechanistic questions further along: if the PTM are old, then they were present before the onset of damage. Resulting in an explanation where the modifications are heterochromatic in origin. However, if the modifications are new, that would imply that the regions have been marked for a particular reason. It could be to signal for repair later on, or it could be to silence the region until repair of the site is possible.

Bibliography

- Alatwi, H.E. & Downs, J.A., 2015. Removal of H2A.Z by INO80 promotes homologous recombination. *EMBO reports*, 16(8), pp.986–994.
- Allfrey, V.G., Faulkner, R. & Mirsky, A.E., 1964. Acetylation and Methylation Of Histones and Their Possible Role In The Regulation Of Rna Synthesis. *Proceedings of the National Academy of Sciences*, 51, pp.786–794.
- Almouzni, G. et al., 1990. Chromatin assembly on replicating DNA in vitro. *Nucleic Acids Research*, 18(19), pp.5767–5774.
- Amendola, P.G. et al., 2017. JMJD-5/KDM8 regulates H3K36me2 and is required for late steps of homologous recombination and genome integrity J. E. Sale, ed. *PLoS Genetics*, 13(2), pp.e1006632–24.
- Arents, G. & Moudrianakis, E.N., 1995. The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. *Proceedings of the National Academy of Sciences*, 92(24), pp.11170–11174.
- Asaithamby, A. & Chen, D.J., 2009. Cellular responses to DNA double-strand breaks after low-dose γ -irradiation. *Nucleic Acids Research*, 37(12), pp.3912–3923.
- Axel, R. et al., 1974. Specific sites of interaction between histones and DNA in chromatin. *Proceedings of the National Academy of Sciences*, 71(10), pp.4101–4105.
- Ayrapetov, M.K. et al., 2014. DNA double-strand breaks promote methylation of histone H3 on lysine 9 and transient formation of repressive chromatin. *Proceedings of the National Academy of Sciences of the United States of America*, 111(25), pp.9169–9174.
- Baldeyron, C. et al., 2011. HP1 α recruitment to DNA damage by p150CAF-1 promotes homologous recombination repair. *The Journal of Cell Biology*, 193(1), pp.81–95.
- Baldock, R.A. et al., 2015. ATM Localization and Heterochromatin Repair Depend on Direct Interaction of the 53BP1-BRCT2 Domain with γ H2AX. *CellReports*, 13(10), pp.2081–2089.
- Bannister, A.J. & Kouzarides, T., 2011. Regulation of chromatin by histone modifications. *Nature Publishing Group*, 21(3), pp.381–395.
- Bargaje, R. et al., 2012. Proximity of H2A.Z containing nucleosome to the transcription start site influences gene expression levels in the mammalian liver and brain. *Nucleic Acids Research*, 40(18), pp.8965–8978.
- Barski, A. et al., 2007. High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell*, 129(4), pp.823–837.

- Bassing, C.H. et al., 2002. Increased ionizing radiation sensitivity and genomic instability in the absence of histone H2AX. *Proceedings of the National Academy of Sciences*, 99(12), pp.8173–8178.
- Begley, C.G. & Ellis, L.M., 2012. Drug development: Raise standards for preclinical cancer research. *Nature*, 483(7391), pp.531–533.
- Beisel, C. & Paro, R., 2011. Silencing chromatin: comparing modes and mechanisms. *Nature Reviews Genetics*, 12(2), pp.123–135.
- Berger, C.E. et al., 2012. p53, a target of estrogen receptor (ER) α , modulates DNA damage-induced growth suppression in ER-positive breast cancer cells. *Journal of Biological Chemistry*, 287(36), pp.30117–30127.
- Berglund, L. et al., 2008. A gene-centric Human Protein Atlas for expression profiles based on antibodies. *Molecular & cellular proteomics : MCP*, 7(10), pp.2019–2027.
- Bernstein, B.E. et al., 2002. Methylation of histone H3 Lys 4 in coding regions of active genes. *Proceedings of the National Academy of Sciences*, 99(13), pp.8695–8700.
- Bird, A.W. et al., 2002. Acetylation of histone H4 by Esa1 is required for DNA double-strand break repair. *Nature*, 419(6905), pp.411–415.
- Black, J.C., Van Rechem, C. & Whetstone, J.R., 2012. Histone Lysine Methylation Dynamics: Establishment, Regulation, and Biological Impact. *Molecular Cell*, 48(4), pp.491–507.
- Blattner, F.R. et al., 1997. The complete genome sequence of Escherichia coli K-12. *Science*, 277(5331), pp.1453–1462.
- Bonaldi, T., Imhof, A. & Regula, J.T., 2004. A combination of different mass spectroscopic techniques for the analysis of dynamic changes of histone modifications. *Proteomics*, 4(5), pp.1382–1396.
- Bonisch, C. & Hake, S.B., 2012. Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Research*, 40(21), pp.10719–10741.
- Bordeaux, J. et al., 2010. Antibody validation. *BioTechniques*, 48(3), pp.197–209.
- Brehove, M. et al., 2015. Histone Core Phosphorylation Regulates DNA Accessibility. *Journal of Biological Chemistry*, 290(37), pp.22612–22621.
- Broering, T.J. et al., 2015. BAZ1B is dispensable for H2AX phosphorylation on Tyrosine 142 during spermatogenesis. *Biology Open*, 4(7), pp.873–884.
- Brownell, J.E. et al., 1996. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell*, 84(6), pp.843–851.
- Buchwalow, I. et al., 2011. Non-specific binding of antibodies in immunohistochemistry: fallacies and facts. *Scientific reports*, 1(1), pp.457–6.

- Caldecott, K.W., 2008. Single-strand break repair and genetic disease. *Nature Reviews Genetics*, 9(8), pp.619–631.
- Campbell, S. et al., 2014. Polycomb repressive complex 2 contributes to DNA double-strand break repair. *Cell cycle (Georgetown, Tex.)*, 12(16), pp.2675–2683.
- Cao, L.-L. et al., 2015. ATM-mediated KDM2A phosphorylation is required for the DNA damage repair. pp.1–13.
- Cao, L.-L., Shen, C. & Zhu, W.-G., 2016. Histone modifications in DNA damage response. *Science China Life Sciences*, 59(3), pp.257–270.
- Caron P, Aymard F, Iacovoni JS, Briois S, Canitrot Y, et al. (2012) Cohesin Protects Genes against γ H2AX Induced by DNA Double-Strand Breaks. *PLOS Genetics* 8(1): e1002460.
- Cea, M. et al., 2016. Evidence for a role of the histone deacetylase SIRT6 in DNA damage response of multiple myeloma cells. *Blood*, 127(9), pp.1138–1150.
- Ceccaldi, R., Rondinelli, B. & D’Andrea, A.D., 2016. Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends in Cell Biology*, 26(1), pp.52–64.
- Celeste, A. et al., 2002. Genomic instability in mice lacking histone H2AX. *Science*, 296(5569), pp.922–927.
- Celeste, A., Difilippantonio, S., et al., 2003. H2AX Haploinsufficiency Modifies Genomic Stability and Tumor Susceptibility. *Cell*, 114(3), pp.371–383.
- Celeste, A., Fernandez-Capetillo, O., et al., 2003. Histone H2AX phosphorylation is dispensable for the initial recognition of DNA breaks. *Nature cell biology*, 5(7), pp.675–679.
- Chapman, J.R., Taylor, M.R.G. & Boulton, S.J., 2012. Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. *Molecular Cell*, 47(4), pp.497–510.
- Chen, H. et al., 2014. MacroH2A1.1 and PARP-1 cooperate to regulate transcription by promoting CBP-mediated H2B acetylation. *Nature structural & molecular biology*, 21(11), pp.981–989.
- Ciccia, A., Constantinou, A. & West, S.C., 2003. Identification and characterization of the human mus81-eme1 endonuclease. *The Journal of biological chemistry*, 278(27), pp.25172–25178.
- Cohen, P.T. & Cohen, P., 1989. Discovery of a protein phosphatase activity encoded in the genome of bacteriophage lambda. Probable identity with open reading frame 221. *The Biochemical journal*, 260(3), pp.931–934.
- Coleman-Derr, D. & Zilberman, D., 2012. Deposition of Histone Variant H2A.Z within Gene Bodies Regulates Responsive Genes T. Kakutani, ed. *PLoS Genetics*, 8(10), pp.e1002988–17.

- Cowell, I.G. et al., 2007. γ H2AX Foci Form Preferentially in Euchromatin after Ionising-Radiation B. Sullivan, ed. *PLoS ONE*, 2(10), pp.e1057–8.
- Curtin, N.J., 2012. DNA repair dysregulation from cancer driver to therapeutic target. *Nature Reviews Cancer*, 12(12), pp.1–17.
- David, S.-A. et al., 2017. An Assessment of Fixed and Native Chromatin Preparation Methods to Study Histone Post-Translational Modifications at a Whole Genome Scale in Skeletal Muscle Tissue. *Biological Procedures Online*, 19(1), pp.1–11.
- Densham, R.M. et al., 2016. Human BRCA1-BARD1 ubiquitin ligase activity counteracts chromatin barriers to DNA resection. *Nature Structural & Molecular Biology*, 23(7), pp.647–655.
- Di Lorenzo, Alessandra and Bedford, Mark T.(2011), Histone arginine methylation, *FEBS Letters*, 585, doi: 10.1016/j.febslet.2010.11.010
- Dirksen, E.H.C. et al., 2006. Human Lymphoblastoid Proteome Analysis Reveals a Role for the Inhibitor of Acetyltransferases Complex in DNA Double-Strand Break Response. *Cancer research*, 66(3), pp.1473–1480.
- Dizdaroglu, M. & Jaruga, P., 2012a. Mechanisms of free radical-induced damage to DNA. *Free Radical Research*, 46(4), pp.382–419.
- Downs, J.A., Lowndes, N.F. & Jackson, S.P., 2000. A role for *Saccharomyces cerevisiae* histone H2A in DNA repair. *Nature*, 408(6815), pp.1001–1004.
- Drake, S.A., 2014. **What are the Energy Range Definitions for the Various Types of Electromagnetic Radiation?**, pp.1–1. Available at: <https://heasarc.gsfc.nasa.gov/docs/heasarc/headates/spectrum.html> [Accessed February 9, 2018].
- Editors, T.T.D., 2008. *LTQ Orbitrap XL*, Thermo Fisher Scientific.
- Egelhofer, T.A. et al., 2010. An assessment of histone-modification antibody quality. *Nature Structural & Molecular Biology*, 18(1), pp.91–93.
- Eirín-López, J.M. et al., 2009. Long-Term Evolution of Histone Families: Old Notions and New Insights into Their Mechanisms of Diversification Across Eukaryotes. In *Evolutionary Biology. Concept, Modelization and Application*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 139–162.
- Etheridge, T.J. et al., 2014. Quantification of DNA-associated proteins inside eukaryotic cells using single-molecule localization microscopy. *Nucleic Acids Research*, 42(19), pp.e146–e146.
- Eyers, C.E. & Gaskell, S., 2014. *Quantitative proteomics*,

- Fan, J.Y. et al., 2002. The essential histone variant H2A.Z regulates the equilibrium between different chromatin conformational states. *Nature Structural Biology*, pp.1–5.
- Fang, J. et al., 2004. Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation. *The Journal of biological chemistry*, 279(51), pp.52812–52815.
- Feng, Q. et al., 2002. Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Current biology : CB*, 12(12), pp.1052–1058.
- Fiedler KL, Cotter RJ. Using glycinylation, a chemical derivatization technique, for the quantitation of ubiquitinated proteins. *Anal Chem*. 2013;85(12):5827–5834. doi:10.1021/ac400398s
- Fnu, S. et al., 2011. Methylation of histone H3 lysine 36 enhances DNA repair by nonhomologous end-joining. *Proceedings of the National Academy of Sciences*, 108(2), pp.540–545.
- Fradet-Turcotte A, Canny MD, Escribano-Díaz C, et al. 53BP1 is a reader of the DNA damage-induced H2A Lys 15 ubiquitin mark. *Nature*. 2013;499(7456):50–54. doi:10.1038/nature12318
- Francia, S. et al., 2012. Site-specific DICER and DROSHA RNA products control the DNA-damage response. *Nature*, 488(7410), pp.231–235.
- Fukagawa, T. et al., 2004. Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nature cell biology*, 6(8), pp.784–791.
- Gallien, S. et al., 2013. Selectivity of LC-MS/MS analysis: implication for proteomics experiments. *Journal of Proteomics*, 81, pp.148–158.
- Gamble, M.J. et al., 2010. The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing. *Genes & development*, 24(1), pp.21–32.
- Gambus, A. et al., 2009. A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome. *The EMBO Journal*, 28(19), pp.2992–3004.
- Garcia, B.A. et al., 2007. Chemical derivatization of histones for facilitated analysis by mass spectrometry. *Nature Protocols*, 2(4), pp.933–938.
- Gehani, S.S. et al., 2010. Polycomb group protein displacement and gene activation through MSK-dependent H3K27me3S28 phosphorylation. *Molecular Cell*, 39(6), pp.886–900.

- Gervais, A.L. & Gaudreau, L., 2009. Discriminating nucleosomes containing histone H2A.Z or H2A based on genetic and epigenetic information. *BMC Molecular Biology*, 10(1), pp.18–9.
- Gieni, R.S. et al., 2014. Polycomb group proteins in the DNA damage response: A link between radiation resistance and “stemness.” *Cell cycle (Georgetown, Tex.)*, 10(6), pp.883–894.
- Gligorovski, S. et al., 2015. Environmental Implications of Hydroxyl Radicals (•OH). *Chem. Rev.*, 115(24), pp.13051–13092.
- Glish, G.L. & Goeringer, D.E., 1984. A tandem quadrupole/time-of-flight instrument for mass spectrometry/mass spectrometry. *Analytical Chemistry*, 56(13), pp.2291–2295.
- Glish, G.L. & Vachet, R.W., 2003. The basics of mass spectrometry in the twenty-first century. *Nature Reviews Drug Discovery*, 2(2), pp.140–150.
- Goldberg, M. et al., 2003. MDC1 is required for the intra-S-phase DNA damage checkpoint. *Nature*, 421(6926), pp.952–956.
- Goldknopf, I.L. et al., 1975. Isolation and characterization of protein A24, a “histone-like” non-histone chromosomal protein. *The Journal of biological chemistry*, 250(18), pp.7182–7187.
- Goldstein, G. et al., 1975. Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proceedings of the National Academy of Sciences*, 72(1), pp.11–15.
- Goodarzi, A.A. et al., 2008. ATM Signaling Facilitates Repair of DNA Double-Strand Breaks Associated with Heterochromatin. *Molecular Cell*, 31(2), pp.167–177.
- Graphodatsky, A.S., Trifonov, V.A. & Stanyon, R., 2011. The genome diversity and karyotype evolution of mammals. *Molecular cytogenetics*, 4(1), p.22.
- Grassi, G., 2003. Inhibitors of DNA methylation and histone deacetylation activate cytomegalovirus promoter-controlled reporter gene expression in human glioblastoma cell line U87. *Carcinogenesis*, 24(10), pp.1625–1635.
- Greaves, J. & Roboz, J., 2013. *Mass Spectrometry for the Novice*, CRC Press.
- Grunstein, M., 1997. Histone acetylation in chromatin structure and transcription. *Nature*, 389(6649), pp.349–352.
- Gupta, D.K. et al., 2016. DNA damage regulation and its role in drug-related phenotypes in the malaria parasites. *Scientific reports*, pp.1–15.
- Gursoy-Yuzugullu, O., Ayrapetov, M.K. & Price, B.D., 2015. Histone chaperone Anp32e removes H2A.Z from DNA double-strand breaks and promotes nucleosome

- reorganization and DNA repair. *Proceedings of the National Academy of Sciences of the United States of America*, 112(24), pp.7507–7512.
- Hamilton C, Hayward RL, Gilbert N. Global chromatin fibre compaction in response to DNA damage. *Biochem Biophys Res Commun*. 2011;414(4):820–825.
doi:10.1016/j.bbrc.2011.10.021
- Harrison, J.C. & Haber, J.E., 2006. Surviving the Breakup: The DNA Damage Checkpoint. *dx.doi.org*, 40(1), pp.209–235.
- Heintzman, N.D. et al., 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genetics*, 39(3), pp.311–318.
- Helt, C.E. et al., 2005. Ataxia telangiectasia mutated (ATM) and ATM and Rad3-related protein exhibit selective target specificities in response to different forms of DNA damage. *The Journal of biological chemistry*, 280(2), pp.1186–1192.
- Henry, R.A. et al., 2016. Quantitative Measurement of Histone Tail Acetylation Reveals Stage-Specific Regulation and Response to Environmental Changes during Drosophila Development. *Biochemistry*, 55(11), pp.1663–1672.
- Heylmann, D. & Kaina, B., 2016. The γ H2AX DNA damage assay from a drop of blood. *Scientific reports*, 6, p.22682.
- Hodawadekar, S.C. & Marmorstein, R., 2007. Chemistry of acetyl transfer by histone modifying enzymes: structure, mechanism and implications for effector design. *Oncogene*, 26(37), pp.5528–5540.
- Huertas, P., 2010. DNA resection in eukaryotes: deciding how to fix the break. *Nature Publishing Group*, 17(1), pp.11–16.
- Hurd, P.J. et al., 2009. Phosphorylation of histone H3 Thr-45 is linked to apoptosis. *The Journal of biological chemistry*, 284(24), pp.16575–16583.
- Ikura, T. et al., 2007. DNA damage-dependent acetylation and ubiquitination of H2AX enhances chromatin dynamics. *Molecular and Cellular Biology*, 27(20), pp.7028–7040.
- Jen, K.-Y. & Cheung, V.G., 2003. Transcriptional response of lymphoblastoid cells to ionizing radiation. *Genome Research*, 13(9), pp.2092–2100.
- Ji, X. et al., 2015. Chromatin proteomic profiling reveals novel proteins associated with histone-marked genomic regions. *Proceedings of the National Academy of Sciences of the United States of America*, pp.201502971–10.
- Jones, B. et al., 2008. The Histone H3K79 Methyltransferase Dot1L Is Essential for Mammalian Development and Heterochromatin Structure W. A. Bickmore, ed. *PLoS Genetics*, 4(9), pp.e1000190–11.

- Josling, G.A. et al., 2012. The Role of Bromodomain Proteins in Regulating Gene Expression. *Genes*, 3(4), pp.320–343.
- Jüttermann, R., Li, E. & Jaenisch, R., 1994. Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. *Proceedings of the National Academy of Sciences*, 91(25), pp.11797–11801.
- Kakarougkas, A. & Jeggo, P.A., 2014. DNA DSB repair pathway choice: an orchestrated handover mechanism. *The British Journal of Radiology*, 87(1035), p.20130685.
- Kanoh, Y., Tamai, K. & Shirahige, K., 2006. Different requirements for the association of ATR–ATRIP and 9-1-1 to the stalled replication forks. *Gene*, 377, pp.88–95.
- Karachentsev, D., 2005. PR-Set7-dependent methylation of histone H4 Lys 20 functions in repression of gene expression and is essential for mitosis. *Genes & Development*, 19(4), pp.431–435.
- Karas, M. & Hillenkamp, F., 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry*, 60(20), pp.2299–2301.
- Kaye, J.A. et al., 2004. DNA breaks promote genomic instability by impeding proper chromosome segregation. *Current biology : CB*, 14(23), pp.2096–2106.
- Kim, Y.-C. et al., 2008. Activation of ATM depends on chromatin interactions occurring before induction of DNA damage. *Nature cell biology*, 11(1), pp.92–96.
- Kleiner, R.E. et al., 2015. Chemical proteomics reveals a γ H2AX-53BP1 interaction in the DNA damage response. *Nature Chemical Biology*, 11(10), pp.1–9.
- Kobayashi, J. et al., 2004. NBS1 and its functional role in the DNA damage response. *DNA Repair*, 3(8-9), pp.855–861.
- Kolas, N.K. et al., 2007. Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. *Science*, 318(5856), pp.1637–1640.
- Kouzarides, T., 2007. Chromatin Modifications and Their Function. *Cell*, 128(4), pp.693–705.
- Krishnan, V., Chow, M. & Wang, Z., 2011. Histone H4 lysine 16 hypoacetylation is associated with defective DNA repair and premature senescence in Zmpste24-deficient mice. In Proceedings of the
- Labib, K. & Hodgson, B., 2007. Replication fork barriers: pausing for a break or stalling for time? *EMBO reports*, 8(4), pp.346–353.
- Lander, E.S. et al., 2001. Initial sequencing and analysis of the human genome. *Nature*, 409(6822), pp.860–921.

- Laskey, R.A. et al., 1978. Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. *Nature*, 275(5679), pp.416–420.
- Lauberth, S.M. et al., 2013. H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell*, 152(5), pp.1021–1036.
- Lee, J.-H. & Paull, T.T., 2007. Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. *Oncogene*, 26(56), pp.7741–7748.
- Levin, S.G., Young, R.W. & Stohler, R.L., 1992. Estimation of median human lethal radiation dose computed from data on occupants of reinforced concrete structures in Nagasaki, Japan. *Health physics*, 63(5), pp.522–531.
- Lewis, K.A. & Wuttke, D.S., 2012. Telomerase and telomere-associated proteins: structural insights into mechanism and evolution. *Structure (London, England : 1993)*, 20(1), pp.28–39.
- Li, A. et al., 2010. Phosphorylation of histone H2A.X by DNA-dependent protein kinase is not affected by core histone acetylation, but it alters nucleosome stability and histone H1 binding. *Journal of Biological Chemistry*, 285(23), pp.17778–17788.
- Lo, W.S. et al., 2000. Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14. *Molecular Cell*, 5(6), pp.917–926.
- Local, A. et al., 2018. Identification of H3K4me1-associated proteins at mammalian enhancers. *Nature Genetics*, 50(1), pp.73–82.
- Lomberk, G., Wallrath, L. & Urrutia, R., 2006. The Heterochromatin Protein 1 family. *Genome Biology*, 7(7), p.228.
- Lou, Z. et al., 2006. MDC1 maintains genomic stability by participating in the amplification of ATM-dependent DNA damage signals. *Molecular Cell*, 21(2), pp.187–200.
- Lu, L. et al., 2015. High-resolution mapping of H4K16 and H3K23 acetylation reveals conserved and unique distribution patterns in Arabidopsis and rice. *Epigenetics*, 10(11), pp.1044–1053.
- Macadangang, B.R. et al., 2014. Evolution of histone 2A for chromatin compaction in eukaryotes. *eLife*, 3, pp.3314–22.
- MacLean, B. et al., 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*, 26(7), pp.966–968.
- Maile, T. et al., 2015. Mass Spectrometric Quantification of Histone Posttranslational Modifications by a Hybrid Chemical Labeling Method. *Molecular & cellular proteomics : MCP*, 14(4), pp.mcp.O114.046573–1158.

- Makarov, A (2000). "Electrostatic axially harmonic orbital trapping: A high-performance technique of mass analysis". *Analytical Chemistry*. 72 (6): 1156–62.
- Mallette, F.E.D.E.R.A. et al., 2012. RNF8- and RNF168-dependent degradation of KDM4A/JMJD2A triggers 53BP1 recruitment to DNA damage sites. *The EMBO Journal*, 31(8), pp.1865–1878.
- Martin, C. & Zhang, Y., 2005. The diverse functions of histone lysine methylation. *Nature Reviews Molecular Cell Biology*, 6(11), pp.838–849.
- Marzluff, W.F. et al., 2002. The Human and Mouse Replication-Dependent Histone Genes. *Genomics*, 80(5), pp.487–498.
- Mattioli, F. et al., 2012. RNF168 Ubiquitinates K13-15 on H2A/H2AX to Drive DNA Damage Signaling. *Cell*, 150(6), pp.1182–1195.
- Metzger, E. et al., 2010. Phosphorylation of histone H3T6 by PKC β (I) controls demethylation at histone H3K4. *Nature*, 464(7289), pp.792–796.
- Michel, M.C., Wieland, T. & Tsujimoto, G., 2009. How reliable are G-protein-coupled receptor antibodies? *Naunyn-Schmiedeberg's Archives of Pharmacology*, 379(4), pp.385–388.
- Mieczkowski, J. et al., 2016. MNase titration reveals differences between nucleosome occupancy and chromatin accessibility. *Nature Communications*, 7, pp.1–11.
- Miller, K.M. et al., 2010. Human HDAC1 and HDAC2 function in the DNA-damage response to promote DNA nonhomologous end-joining. *Nature structural & molecular biology*, 17(99), pp.1144–1151.
- Mogal, A. & Abdulkadir, S.A., 2006. Effects of Histone Deacetylase Inhibitor (HDACi); Trichostatin-A (TSA) on the expression of housekeeping genes. *Molecular and Cellular Probes*, 20(2), pp.81–86.
- Mole, R.H., 1984. The LD50 for uniform low LET irradiation of man. *The British Journal of Radiology*, 57(677), pp.355–369.
- Moritz, B., Becker, P.B. & Göpfert, U., 2015. CMV promoter mutants with a reduced propensity to productivity loss in CHO cells. *Scientific reports*, 5(1), pp.917–8.
- Mosammaparast, N., Kim, H. & Laurent, B., 2013. The histone demethylase LSD1/KDM1A promotes the DNA damage response. *The Journal of cell ...*, 203(3), pp.457–470.
- Murray, J.M., Stiff, T. & Jeggo, P.A., 2012. DNA double-strand break repair within heterochromatic regions. *Biochemical Society transactions*, 40(1), pp.173–178.
- Nakamura, T.M. et al., 2004. Histone H2A Phosphorylation Controls Rhp9 Recruitment at DNA Breaks, Maintains Checkpoint Arrest, and Influences DNA Repair in Fission Yeast. *Molecular and Cellular Biology*, 24(14), pp.6215–6230.

- Nashun, B. et al., 2015. Continuous Histone Replacement by Hira Is Essential for Normal Transcriptional Regulation and De Novo DNA Methylation during Mouse Oogenesis. *Molecular Cell*, 60(4), pp.611–625.
- Nelson, T., Hsieh, T.S. & Brutlag, D., 1979. Extracts of Drosophila embryos mediate chromatin assembly in vitro. *Proceedings of the National Academy of Sciences*, 76(11), pp.5510–5514.
- Neumayer, G. & Nguyen, M.D., 2014. TPX2 Impacts Acetylation of Histone H4 at Lysine 16: Implications for DNA Damage Response W.-G. Zhu, ed. *PLoS ONE*, 9(11), pp.e110994–9.
- Olcina, M.M. & Hammond, E.M., 2013. Hypoxia and the DNA Damage Response. In *Hypoxia and Cancer*. Cancer Drug Discovery and Development. New York, NY: Springer New York, pp. 21–41.
- Oudet, P., Gross-Bellard, M. & Chambon, P., 1975. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell*, 4(4), pp.281–300.
- Owen DJ, et al. (2000) The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. *EMBO J* 19(22):6141-9
- Ozaki, T. et al., 2003. Chromosomal alterations in osteosarcoma cell lines revealed by comparative genomic hybridization and multicolor karyotyping. *Cancer genetics and cytogenetics*, 140(2), pp.145–152.
- Ozdemir, A. et al., 2005. Characterization of lysine 56 of histone H3 as an acetylation site in Saccharomyces cerevisiae. *The Journal of biological chemistry*, 280(28), pp.25949–25952.
- Pan, M.-R. et al., 2011. Monoubiquitination of H2AX protein regulates DNA damage response signaling. *Journal of Biological Chemistry*, 286(32), pp.28599–28607.
- Papazyan, R. et al., 2014. Methylation of histone H3K23 blocks DNA damage in pericentric heterochromatin during meiosis. *eLife*, 3, pp.153–23.
- Parthun, M.R., 2007. Hat1: the emerging cellular roles of a type B histone acetyltransferase. *Oncogene*, 26(37), pp.5319–5328.
- Paulovich, A.G. et al., 2010. Interlaboratory study characterizing a yeast performance standard for benchmarking LC-MS platform performance. *Molecular & cellular proteomics : MCP*, 9(2), pp.242–254.
- Pchelintsev, N.A., Adams, P.D. & Nelson, D.M., 2016. Critical Parameters for Efficient Sonication and Improved Chromatin Immunoprecipitation of High Molecular Weight Proteins M. Wu, ed. *PLoS ONE*, 11(1), pp.e0148023–11.

- Pei, H., Zhang, L., Luo, K., Qin, Y., Chesi, M., Fei, F., Bergsagel, P.L., Wang, L., You, Z. & Lou, Z., 2011a. MMSET regulates histone H4K20 methylation and 53BP1 accumulation at DNA damage sites. *Nature*, 470(7332), pp.124–128.
- Pellegrino S, Michelena J, Teloni F, Imhof R, Altmeyer M. Replication-Coupled Dilution of H4K20me2 Guides 53BP1 to Pre-replicative Chromatin. *Cell Rep*. 2017;19(9):1819–1831. doi:10.1016/j.celrep.2017.05.016
- Pepys, M.B., 2006. Amyloidosis. *Annual Review of Medicine*, 57(1), pp.223–241.
- Pesavento, J.J. et al., 2007. Certain and Progressive Methylation of Histone H4 at Lysine 20 during the Cell Cycle. *Molecular and Cellular Biology*, 28(1), pp.468–486.
- Petersen, S. et al., 2001. AID is required to initiate Nbs1/gamma-H2AX focus formation and mutations at sites of class switching. *Nature*, 414(6864), pp.660–665.
- Pfister, S.X. et al., 2014. SETD2-Dependent Histone H3K36 Trimethylation Is Required for Homologous Recombination Repair and Genome Stability. *CellReports*, 7(6), pp.2006–2018.
- Pineda, G. et al., 2015. Proteomics studies of the interactome of RNA polymerase II C-terminal repeated domain. *BMC Research Notes*, 8(1), pp.1–22.
- Podhorecka, M., Skladanowski, A. & Bozko, P., 2010. H2AX Phosphorylation: Its Role in DNA Damage Response and Cancer Therapy. *Journal of nucleic acids*, 2010(5), pp.1–9.
- Prassas, I. & Diamandis, E.P., 2014. Translational researchers beware! Unreliable commercial immunoassays (ELISAs) can jeopardize your research. *Clinical Chemistry and Laboratory Medicine (CCLM)*, 52(6), pp.1–2.
- Prise, K.M., 2017. Ionizing Radiation Therapy. In M. Schwab, ed. *Encyclopedia of Cancer*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 3150–3150.
- A.H.Q. (2014). Practical considerations and current limitations in quantitative mass spectrometry-based proteomics. In *Quantitative proteomics*. pubs.rsc.org
- Quintana-Murci, L. & Fellous, M., 2001. The Human Y Chromosome: The Biological Role of a “Functional Wasteland.” *Journal of biomedicine & biotechnology*, 1(1), pp.18–24.
- Radman, M., 2016. Protein damage, radiation sensitivity and aging. *DNA Repair*, 44, pp.186–192.
- Ran, F.A. et al., 2013. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11), pp.2281–2308.
- Rappsilber, J., Mann, M. & Ishihama, Y., 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols*, 2(8), pp.1896–1906.

- Rea, S. et al., 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*, 406(6796), pp.593–599.
- Riballo, E. et al., 2004. A pathway of double-strand break rejoining dependent upon ATM, Artemis, and proteins locating to gamma-H2AX foci. *Molecular Cell*, 16(5), pp.715–724.
- Rogakou, E.P. & Bonner, W.M., 2000. A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Current biology : CB*, 10(15), pp.886–895.
- Rogakou, E.P. et al., 1998. DNA Double-stranded Breaks Induce Histone H2AX Phosphorylation on Serine 139. *The Journal of biological chemistry*, 273(10), pp.5858–5868.
- Rudnick, P.A. et al., 2010. Performance metrics for liquid chromatography-tandem mass spectrometry systems in proteomics analyses. *Molecular & cellular proteomics : MCP*, 9(2), pp.225–241.
- Saksouk, N., Simboeck, E. & DEjardin, J., 2015. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics & Chromatin*, 8(1), p.3.
- Sandman, K. & Reeve, J.N., 2000. Structure and functional relationships of archaeal and eukaryal histones and nucleosomes. *Archives of microbiology*, 173(3), pp.165–169.
- Santos-Rosa, H. et al., 2002. Active genes are tri-methylated at K4 of histone H3. *Nature*, 419(6905), pp.407–411.
- Sechi, S., 2007. *Quantitative Proteomics by Mass Spectrometry*, Springer Science & Business Media.
- Sharma, A.K. et al., 2015. Dynamic alteration in H3 serine 10 phosphorylation is G1-phase specific during ionization radiation induced DNA damage response in human cells. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, 773, pp.83–91.
- Sharma, G.G. et al., 2010. MOF and Histone H4 Acetylation at Lysine 16 Are Critical for DNA Damage Response and Double-Strand Break Repair. *Molecular and Cellular Biology*, 30(14), pp.3582–3595.
- Shema E, Jones D, Shores N, Donohue L, Ram O, Bernstein BE. Single-molecule decoding of combinatorially modified nucleosomes. *Science*. 2016;352(6286):717–721. doi:10.1126/science.aad7701
- Shi, X. et al., 2006. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature*, 442(7098), pp.96–99.
- Shi, Y. et al., 2004. Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1. *Cell*, 119(7), pp.941–953.

- Shumaker, D.K. et al., 2006. Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proceedings of the National Academy of Sciences of the United States of America*, 103(23), pp.8703–8708.
- Sidoli, S. et al., 2015. Drawbacks in the use of unconventional hydrophobic anhydrides for histone derivatization in bottom-up proteomics PTM analysis. *Proteomics*, 15(9), pp.1459–1469.
- Sie, L., Loong, S. & Tan, E.K., 2009. Utility of lymphoblastoid cell lines. *Journal of Neuroscience Research*, 87(9), pp.1953–1959.
- Sinha, R.P. & Hader, D.P., 2002. UV-induced DNA damage and repair: a review. *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology*, 1(4), pp.225–236.
- Smith, K.D., Fu, M.A. & Brown, E.J., 2009. Tim–Tipin dysfunction creates an indispensable reliance on the ATR–Chk1 pathway for continued DNA synthesis. *The Journal of Cell Biology*, 187(1), pp.15–23.
- Soldi, M. , Cuomo, A. and Bonaldi, T. (2014), Improved bottom-up strategy to efficiently separate hypermodified histone peptides through ultra-HPLC separation on a bench top Orbitrap instrument. *Proteomics*, 14: 2212-2225.
- Solier, S. et al., 2009. Death receptor-induced activation of the Chk2- and histone H2AX-associated DNA damage response pathways. *Molecular and Cellular Biology*, 29(1), pp.68–82.
- Sonntag, von, C., 2006. *Free-radical-induced DNA damage and its repair*,
- Soria, G. & Almouzni, G., 2013. Differential contribution of HP1 proteins to DNA end resection and homology-directed repair. *Cell cycle (Georgetown, Tex.)*, 12(3), pp.422–429.
- Stathopoulos, P.B. et al., 2008. Sonication of proteins causes formation of aggregates that resemble amyloid. *Protein Science*, 13(11), pp.3017–3027.
- Steen, H. & Mann, M., 2004. The ABC“s (and XYZ”s) of peptide sequencing. *Nature Reviews Molecular Cell Biology*, 5(9), pp.699–711.
- Stewart, G.S. et al., 2003. MDC1 is a mediator of the mammalian DNA damage checkpoint. *Nature*, 421(6926), pp.961–966.
- Strahl, B.D. & Allis, C.D., 2000. The language of covalent histone modifications. *Nature*, 403(6765), pp.41–45.
- Stucki, M. et al., 2005. MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks. *Cell*, 123(7), pp.1213–1226.

- Sun, L., Zhu, G. & Dovichi, N.J., 2012. Comparison of the LTQ-Orbitrap Velos and the Q-Exactive for proteomic analysis of 1-1000ng RAW 264.7 cell lysate digests. *Rapid Communications in Mass Spectrometry*, 27(1), pp.157–162.
- Sun, Y. et al., 2005. A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM. *Proceedings of the National Academy of Sciences*, 102(37), pp.13182–13187.
- Sun, Y. et al., 2009. Histone H3 methylation links DNA damage detection to activation of the tumour suppressor Tip60. *Nature cell ...*, 11(11), pp.1376–1382.
- Sweet, S.M.M. et al., 2010. Kinetics of Re-establishing H3K79 Methylation Marks in Global Human Chromatin. *The Journal of biological chemistry*, 285(43), pp.32778–32786.
- Tabb, D.L. et al., 2010. Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *Journal of Proteome Research*, 9(2), pp.761–776.
- Talbert, P.B. & Henikoff, S., 2010. Histone variants — ancient wrap artists of the epigenome. *Nature Reviews Molecular Cell Biology*, 11(4), pp.264–275.
- Tamburini, B.A. & Tyler, J.K., 2005. Localized Histone Acetylation and Deacetylation Triggered by the Homologous Recombination Pathway of Double-Strand DNA Repair. *Molecular and Cellular Biology*, 25(12), pp.4903–4913.
- Tang, J. et al., 2013. Acetylation limits 53BP1 association with damaged chromatin to promote homologous recombination. *Nature structural & molecular biology*, 20(3), pp.317–325.
- Taverna, S.D. et al., 2007. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nature Structural & Molecular Biology*, 14(11), pp.1025–1040.
- The UniProt Consortium, 2017. UniProt: the universal protein knowledgebase. *Nucleic Acids Research*, 45(D1), pp.D158–D169.
- Tjeertes, J.V., Miller, K.M. & Jackson, S.P., 2009. Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells. *The EMBO Journal*, 28(13), pp.1878–1889.
- Tong, Z. et al., 2016. SIRT7 Is Activated by DNA and Deacetylates Histone H3 in the Chromatin Context. *ACS Chemical Biology*, 11(3), pp.742–747.
- Travers, A. & Muskhelishvili, G., 2015. DNA structure and function. *FEBS Journal*, 282(12), pp.2279–2295.
- Tryndyak, V.P., Kovalchuk, O. & Pogribny, I.P., 2006. Loss of DNA methylation and histone H4 lysine 20 trimethylation in human breast cancer cells is associated with aberrant expression of DNA methyltransferase 1, Suv4-20h2 histone

- methyltransferase and methyl-binding proteins. *Cancer biology & therapy*, 5(1), pp.65–70.
- Tsai, W.-W. et al., 2010. TRIM24 links a non-canonical histone signature to breast cancer. *Nature*, 468(7326), pp.927–932.
- Tseng, Z. et al., 2014. Using Native Chromatin Immunoprecipitation to Interrogate Histone Variant Protein Deposition in Embryonic Stem Cells. In R. H. Morse, ed. *Chromatin Remodeling*. Methods in Molecular Biology. New York, NY: Springer New York, pp. 11–22.
- Tsukada, Y.-I. et al., 2006. Histone demethylation by a family of JmjC domain-containing proteins. *Nature*, 439(7078), pp.811–816.
- Turner, B., 2001. ChIP with Native Chromatin: Advantages and Problems Relative to Methods Using Cross-Linked Material.
- Unnikrishnan, A., Gafken, P.R. & Tsukiyama, T., 2010. Dynamic changes in histone acetylation regulate origins of DNA replication. *Nature structural & molecular biology*, 17(4), pp.430–437.
- Vallur, A.C. et al., 2002. Effects of Hydrogen Bonding within a Damaged Base Pair on the Activity of Wild Type and DNA-intercalating Mutants of Human Alkyladenine DNA Glycosylase. *The Journal of biological chemistry*, 277(35), pp.31673–31678.
- Vazquez, B.N. et al., 2016. SIRT7 promotes genome integrity and modulates non-homologous end joining DNA repair. *The EMBO Journal*, 35(14), pp.1488–1503.
- Verdaasdonk, J.S. & Bloom, K., 2011. Centromeres: unique chromatin structures that drive chromosome segregation. *Nature Reviews Molecular Cell Biology*, 12(5), pp.320–332.
- Vilenchik, M.M. & Knudson, A.G., 2003. Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer. *Proceedings of the National Academy of Sciences*, 100(22), pp.12871–12876.
- Voigt, P., Tee, W.-W. & Reinberg, D., 2013. A double take on bivalent promoters. *Genes & Development*, 27(12), pp.1318–1338.
- Vondriska, T. M. and Jensen, O. N. (2014), Proteomics in Chromatin Biology and Epigenetics. *Proteomics*, 14: 2105-2108
- Wakeman, T.P. et al., 2012. Bat3 facilitates H3K79 dimethylation by DOT1L and promotes DNA damage-induced 53BP1 foci at G1/G2 cell-cycle phases. *The EMBO Journal*, 31(9), pp.2169–2181.
- Wang, B. & Elledge, S.J., 2007. Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brca1/Brcc36 complex in response to DNA damage. *Proceedings of the National Academy of Sciences of the United States of America*, 104(52), pp.20759–20763.

- Wang, H. et al., 2004. Role of histone H2A ubiquitination in Polycomb silencing. *Nature*, 431(7010), pp.862–868.
- Ward, I.M. et al., 2003. Accumulation of checkpoint protein 53BP1 at DNA breaks involves its binding to phosphorylated histone H2AX. *The Journal of biological chemistry*, 278(22), pp.19579–19582.
- Waters Corporation, 2016. *Waters SYNAPT G2-S HDMS Mass Spectrometer - Overview and Maintenance Guide*, Waters Corporation.
- Watson, J.D. & Crick, F.H., 1953. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171(4356), pp.737–738.
- West, S.M. et al., 2012. Electrostatic Interactions between Arginines and the Minor Groove in the Nucleosome. *Journal of Biomolecular Structure and Dynamics*, 27(6), pp.861–866.
- Woo, Y.H. & Li, W.H., 2012. Evolutionary Conservation of Histone Modifications in Mammals. *Molecular Biology and Evolution*, 29(7), pp.1757–1767.
- World Health Organisation, 2016. WHO | Health effects of the Chernobyl accident: an overview. pp.1–8. Available at: http://www.who.int/ionizing_radiation/chernobyl/background/en/ [Accessed May 23, 2018].
- Wu, T. et al., 2014. Histone variant H2A.X deposition pattern serves as a functional epigenetic mark for distinguishing the developmental potentials of iPSCs. *Cell stem cell*, 15(3), pp.281–294.
- Xiao, A. et al., 2008. WSTF regulates the H2A.X DNA damage response via a novel tyrosine kinase activity. *Nature*, 457(7225), pp.57–64.
- Xu, G. & Jaffrey, S.R., 2013. Proteomic identification of protein ubiquitination events. *Biotechnology and Genetic Engineering Reviews*, 29(1), pp.73–109.
- Xu, Y. et al., 2012. Histone H2A.Z Controls a Critical Chromatin Remodeling Step Required for DNA Double-Strand Break Repair. *Molecular Cell*, 48(5), pp.723–733.
- Yalow, R.S. & Berson, S.A., 2007. Immunoassay of Endogenous Plasma Insuline in Man. *The Journal of Clinical Investigation*, 39(7), pp.1157–1175.
- Yamada, S. et al., 2017. The histone variant H2A.Z promotes initiation of meiotic recombination in fission yeast. *Nucleic Acids Research*, 46(2), pp.609–620.
- Yamashita, M. & Fenn, J.B., 1983. Electrospray Ion Source. Another Variation on the Free-Jet Theme. *The Journal of Physical Chemistry*, 88(20), pp.4451–4459.
- Yan, J.C.A.Q., 2012. Histone ubiquitination and deubiquitination in transcription, DNA damage response, and cancer. pp.1–9.

- Yang, X. et al., 2013. Histone Acetyltransferase 1 Promotes Homologous Recombination in DNA Repair by Facilitating Histone Turnover. *The Journal of biological chemistry*, 288(25), pp.18271–18282.
- Yuan, J. et al., 2014. Histone H3-K56 acetylation is important for genomic stability in mammals. *Cell cycle (Georgetown, Tex.)*, 8(11), pp.1747–1753.
- Zhang, L. et al., 2003. Identification of novel histone post-translational modifications by peptide mass fingerprinting. *Chromosoma*, 112(2), pp.77–86.
- Zhang, P.-Y. et al., 2016. Dicer interacts with SIRT7 and regulates H3K18 deacetylation in response to DNA damaging agents. *Nucleic Acids Research*, 44(8), pp.3629–3642.
- Zhou, L. et al., 2016. Evidence that ubiquitylated H2B corrals hDot1L on the nucleosomal surface to induce H3K79 methylation. *Nature Communications*, 7, pp.1–9.
- Zhu, Q. & Wani, A.A., 2010. Histone modifications: crucial elements for damage response and chromatin restoration. *Journal of Cellular Physiology*, 281, pp.n/a–n/a.
- Zhuo, S. et al., 1993. Expression, purification, crystallization, and biochemical characterization of a recombinant protein phosphatase. *The Journal of biological chemistry*, 268(24), pp.17754–17761.

Appendices

Appendix I: Table of enriched phosphorylated non-histone peptides. (Chapter 3)

Protein Description	Peptide Sequence
28 kDa heat- and acid-stable phosphoprotein OS=Homo sapiens GN=PDAP1 PE=1 SV=1	SLDSDESEDEEDDYQQK
28 kDa heat- and acid-stable phosphoprotein OS=Homo sapiens GN=PDAP1 PE=1 SV=1	SLDSDESEDEEDDYQQK
40S ribosomal protein S3 OS=Homo sapiens GN=RPS3 PE=1 SV=2	DEILPTTPISEQK
40S ribosomal protein S6 OS=Homo sapiens GN=RPS6 PE=1 SV=1	RLSSLR
40S ribosomal protein S6 OS=Homo sapiens GN=RPS6 PE=1 SV=1	RLSSLR
60S acidic ribosomal protein P2 OS=Homo sapiens GN=RPLP2 PE=1 SV=1	KEESEESDDDMGFGLFD
60S acidic ribosomal protein P2 OS=Homo sapiens GN=RPLP2 PE=1 SV=1	KEESEESDDDMGFGLFD
60S ribosomal protein L12 OS=Homo sapiens GN=RPL12 PE=1 SV=1	IGPLGLSPK
Alpha-adducin OS=Homo sapiens GN=ADD1 PE=1 SV=2	SPGSPVGEGTGSPPK
Apoptosis inhibitor 5 OS=Homo sapiens GN=API5 PE=1 SV=3	ASEDTTSGSPPKK
Apoptosis inhibitor 5 OS=Homo sapiens GN=API5 PE=1 SV=3	ASEDTTSGSPPKK
Apoptotic chromatin condensation inducer in the nucleus OS=Homo sapiens GN=ACIN1 PE=1 SV=2	LSEGSQPAEEEEEDQETPSR
Apoptotic chromatin condensation inducer in the nucleus OS=Homo sapiens GN=ACIN1 PE=1 SV=2	KISVVSATK
Apoptotic chromatin condensation inducer in the nucleus OS=Homo sapiens GN=ACIN1 PE=1 SV=2	RLSQPESAEK
Arginine/serine-rich coiled-coil protein 2 OS=Homo sapiens GN=RSRC2 PE=1 SV=1	EQSEVSVSPR
Bcl-2-associated transcription factor 1 OS=Homo sapiens GN=BCLAF1 PE=1 SV=2	FNDSEGDDTEETEDYR
Bcl-2-associated transcription factor 1 OS=Homo sapiens GN=BCLAF1 PE=1 SV=2	DLFDYSPPLHK
Bcl-2-associated transcription factor 1 OS=Homo sapiens GN=BCLAF1 PE=1 SV=2	IDISPSTLR
Bcl-2-associated transcription factor 1 OS=Homo sapiens GN=BCLAF1 PE=1 SV=2	LKDLFDYSPPLHK
Bcl-2-associated transcription factor 1 OS=Homo sapiens GN=BCLAF1 PE=1 SV=2	SPEIHR

Bromodomain-containing protein 3 OS=Homo sapiens GN=BRD3 PE=1 SV=1	SESPPPLSDPK
Centrin-2 OS=Homo sapiens GN=CETN2 PE=1 SV=1	RMSPKPELTEEQK
Chromatin complexes subunit BAP18 OS=Homo sapiens GN=BAP18 PE=1 SV=1	VYEDSGIPLPAESPK
Chromobox protein homolog 1 OS=Homo sapiens GN=CBX1 PE=1 SV=1	KADSDSEDKGEEKPK
Chromobox protein homolog 1 OS=Homo sapiens GN=CBX1 PE=1 SV=1	KADSDSEDKGEEKPK
Chromobox protein homolog 1 OS=Homo sapiens GN=CBX1 PE=1 SV=1	KADSDSEDKGEEKPK
Chromosome alignment-maintaining phosphoprotein 1 OS=Homo sapiens GN=CHAMP1 PE=1 SV=2	KPGPPLSPEIR
Coiled-coil domain-containing protein 137 OS=Homo sapiens GN=CCDC137 PE=1 SV=1	VQAGPGSPR
Coiled-coil domain-containing protein 86 OS=Homo sapiens GN=CCDC86 PE=1 SV=1	AGLGSPERPPK
Coiled-coil domain-containing protein 86 OS=Homo sapiens GN=CCDC86 PE=1 SV=1	QPEYSPESPR
Coiled-coil domain-containing protein 86 OS=Homo sapiens GN=CCDC86 PE=1 SV=1	LGGLRPESPESLTSVSR
Cyclic AMP-dependent transcription factor ATF-2 OS=Homo sapiens GN=ATF2 PE=1 SV=4	MPLDLSPLATPIIR
Cyclin-dependent kinase 3 OS=Homo sapiens GN=CDK3 PE=1 SV=1	IGEGTYGVVYK
DNA damage-binding protein 2 OS=Homo sapiens GN=DDB2 PE=1 SV=1	SRSPLELEPEAK
DNA mismatch repair protein Msh6 OS=Homo sapiens GN=MSH6 PE=1 SV=2	SEEDNEIESEEEVQPK
DNA repair protein complementing XP-C cells OS=Homo sapiens GN=XPC PE=1 SV=4	VIKDEALSDGDDLRL
DNA topoisomerase 2-beta OS=Homo sapiens GN=TOP2B PE=1 SV=3	FDSNEEDSASVFSPSFGK
DNA topoisomerase 2-beta OS=Homo sapiens GN=TOP2B PE=1 SV=3	VVEAVNSDSDSEFGIPK
Drebrin OS=Homo sapiens GN=DBN1 PE=1 SV=4	SPSDSSTASTPVAEQIER
E3 ubiquitin-protein ligase TRIP12 OS=Homo sapiens GN=TRIP12 PE=1 SV=1	SESPPAELPSLR
Elongation factor 1-beta OS=Homo sapiens GN=EEF1B2 PE=1 SV=3	DDDDIDLFGSDDEEESEAKR
Elongation factor 1-beta OS=Homo sapiens GN=EEF1B2 PE=1 SV=3	DDDDIDLFGSDDEEESEAK
Endoplasmic reticulum protein OS=Homo sapiens GN=HSP90B1 PE=1 SV=1	EESDDEAAVEEEEEEEK
Eukaryotic translation initiation factor 5B OS=Homo sapiens GN=EIF5B PE=1 SV=4	KQSFDDNDSEELEDK
Eukaryotic translation initiation factor 5B OS=Homo sapiens GN=EIF5B PE=1 SV=4	NKPGPNIESGNEDDDASFK

Forkhead box protein K2 OS=Homo sapiens GN=FO XK2 PE=1 SV=3	EGSPAPLEPEPGAAQPK
Fos-related antigen 2 OS=Homo sapiens GN=FOSL2 PE=1 SV=1	RSPPAPGLQPMR
Fos-related antigen 2 OS=Homo sapiens GN=FOSL2 PE=1 SV=1	SPPAPGLQPMR
Heat shock protein beta-1 OS=Homo sapiens GN=HSPB1 PE=1 SV=2	QLSSGVSEIR
Heat shock protein HSP 90-alpha OS=Homo sapiens GN=HSP90AA1 PE=1 SV=5	ESEDKPEIEDVGSDEEEEEK
Heat shock protein HSP 90-alpha OS=Homo sapiens GN=HSP90AA1 PE=1 SV=5	ESEDKPEIEDVGSDEEEEEKK
Heat shock protein HSP 90-beta OS=Homo sapiens GN=HSP90AB1 PE=1 SV=4	IEDVGSDEEDDSGK
Heat shock protein HSP 90-beta OS=Homo sapiens GN=HSP90AB1 PE=1 SV=4	EKEISDDEAEEEEKGEK
Heat shock protein HSP 90-beta OS=Homo sapiens GN=HSP90AB1 PE=1 SV=4	IEDVGSDEEDDSGKDK
Heat shock protein HSP 90-beta OS=Homo sapiens GN=HSP90AB1 PE=1 SV=4	IEDVGSDEEDDSGKDK
Hepatoma-derived growth factor OS=Homo sapiens GN=HDGF PE=1 SV=1	AGDLLEDSPK
Hepatoma-derived growth factor OS=Homo sapiens GN=HDGF PE=1 SV=1	RAGDLLEDSPK
Heterogeneous nuclear ribonucleoprotein A1 OS=Homo sapiens GN=HNRNPA1 PE=1 SV=5	SESPKEPEQLR
Heterogeneous nuclear ribonucleoprotein D0 OS=Homo sapiens GN=HNRNPD PE=1 SV=1	NEEDEGHNSNSPR
Heterogeneous nuclear ribonucleoprotein D0 OS=Homo sapiens GN=HNRNPD PE=1 SV=1	NEEDEGHNSNSPR
Heterogeneous nuclear ribonucleoprotein K OS=Homo sapiens GN=HNRNPK PE=1 SV=1	DYDDMSPR
Heterogeneous nuclear ribonucleoprotein U OS=Homo sapiens GN=HNRNPU PE=1 SV=6	AKSPQPPVEEEDHFDDTVVCLDTYNCDLHF K
Heterogeneous nuclear ribonucleoproteins C1/C2 OS=Homo sapiens GN=HNRNPC PE=1 SV=4	NDKSEEEQSSSVK
Heterogeneous nuclear ribonucleoproteins C1/C2 OS=Homo sapiens GN=HNRNPC PE=1 SV=4	EAEEGEDDRDSANGEDDS
Heterogeneous nuclear ribonucleoproteins C1/C2 OS=Homo sapiens GN=HNRNPC PE=1 SV=4	VSGNTSR
High mobility group nucleosome-binding domain-containing protein 3 OS=Homo sapiens GN=HMGN3 PE=1 SV=2	RKSPENTEGK
High mobility group protein HMG-I/HMG-Y OS=Homo sapiens GN=HMGA1 PE=1 SV=3	EEEEGISQESSEEEQ
High mobility group protein HMG-I/HMG-Y OS=Homo sapiens GN=HMGA1 PE=1 SV=3	KQPPVSPGTALVGSQK

High mobility group protein HMG-I/HMG-Y OS=Homo sapiens GN=HMGA1 PE=1 SV=3	EEEEGISQESSEEEQ
High mobility group protein HMG-I/HMG-Y OS=Homo sapiens GN=HMGA1 PE=1 SV=3	SSQPLASK
High mobility group protein HMG-I/HMG-Y OS=Homo sapiens GN=HMGA1 PE=1 SV=3	EEEEGISQESSEEEQ
High mobility group protein HMG-I/HMG-Y OS=Homo sapiens GN=HMGA1 PE=1 SV=3	KQPPVSPGTALVGSQK
High mobility group protein HMG-I/HMG-Y OS=Homo sapiens GN=HMGA1 PE=1 SV=3	QPPVSPGTALVGSQK
High mobility group protein HMGI-C OS=Homo sapiens GN=HMGA2 PE=1 SV=1	KPAQEETEETSSQESAED
High mobility group protein HMGI-C OS=Homo sapiens GN=HMGA2 PE=1 SV=1	QQQEPTGEPSPK
High mobility group protein HMGI-C OS=Homo sapiens GN=HMGA2 PE=1 SV=1	KQQQEPTGEPSPK
HIRA-interacting protein 3 OS=Homo sapiens GN=HIRIP3 PE=1 SV=3	ESEQESEEIILAQK
Histone H1.0 OS=Homo sapiens GN=H1F0 PE=1 SV=3	GVGASGSFR
Histone H1.2 OS=Homo sapiens GN=HIST1H1C PE=1 SV=2	KAAGGATPK
Histone H1.4 OS=Homo sapiens GN=HIST1H1E PE=1 SV=2	GTGASGSFK
Histone H1.4 OS=Homo sapiens GN=HIST1H1E PE=1 SV=2	SETAPAAPAAPAPAEKTPVK
Histone H1.4 OS=Homo sapiens GN=HIST1H1E PE=1 SV=2	KAPKSPAK
Histone H1.4 OS=Homo sapiens GN=HIST1H1E PE=1 SV=2	KASGPPVSELITK
Histone H1.4 OS=Homo sapiens GN=HIST1H1E PE=1 SV=2	ASGPPVSELITK
Histone H1.4 OS=Homo sapiens GN=HIST1H1E PE=1 SV=2	KATGAATPK
Histone H1.4 OS=Homo sapiens GN=HIST1H1E PE=1 SV=2	APKSPAK
Histone H1x OS=Homo sapiens GN=H1FX PE=1 SV=1	AGGSAALSPSK
Histone H1x OS=Homo sapiens GN=H1FX PE=1 SV=1	AGGSAALSPSKK
Histone H2AX OS=Homo sapiens GN=H2AFX PE=1 SV=2	KATQASQEY
Histone H2AX OS=Homo sapiens GN=H2AFX PE=1 SV=2	ATQASQEY
Histone H3.2 OS=Homo sapiens GN=HIST2H3A PE=1 SV=3	KSTGGKAPR
Histone H3.2 OS=Homo sapiens GN=HIST2H3A PE=1 SV=3	KSTGGKAPR
Histone H3.2 OS=Homo sapiens GN=HIST2H3A PE=1 SV=3	STELLIR

Histone H4 OS=Homo sapiens GN=HIST1H4A PE=1 SV=2	RISGLIYEETR
Interferon-inducible double-stranded RNA-dependent protein kinase activator A OS=Homo sapiens GN=PRKRA PE=1 SV=1	EDSGTFSLGK
Interleukin enhancer-binding factor 3 OS=Homo sapiens GN=ILF3 PE=1 SV=3	RPMEEDGEEKSPSK
Interleukin enhancer-binding factor 3 OS=Homo sapiens GN=ILF3 PE=1 SV=3	RPMEEDGEEKSPSK
LIM domain and actin-binding protein 1 OS=Homo sapiens GN=LIMA1 PE=1 SV=1	ASSLSESPPK
LIM domain and actin-binding protein 1 OS=Homo sapiens GN=LIMA1 PE=1 SV=1	ETPHSPGVEDAPIAK
LIM domain and actin-binding protein 1 OS=Homo sapiens GN=LIMA1 PE=1 SV=1	ETPHSPGVEDAPIAK
MARCKS-related protein OS=Homo sapiens GN=MARCKSL1 PE=1 SV=2	GDVTAEAAAGASPAK
Methyl-CpG-binding protein 2 OS=Homo sapiens GN=MECP2 PE=1 SV=1	AETSESGSGAPAVPEASAPK
Neuroblast differentiation-associated protein AHNAK OS=Homo sapiens GN=AHNAK PE=1 SV=2	GGVTGSPEASISGSK
Neuroblast differentiation-associated protein AHNAK OS=Homo sapiens GN=AHNAK PE=1 SV=2	GHYEVTGSDDETGK
Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1 OS=Homo sapiens GN=NUCKS1 PE=1 SV=1	NSQEDSEDSKDVK
Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1 OS=Homo sapiens GN=NUCKS1 PE=1 SV=1	VVDYSQFQESDDADEDYGR
Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1 OS=Homo sapiens GN=NUCKS1 PE=1 SV=1	DDSHAEDSEDEKEDHK
Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1 OS=Homo sapiens GN=NUCKS1 PE=1 SV=1	DSGSDDEFLMEDDDSDYGSSK
Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1 OS=Homo sapiens GN=NUCKS1 PE=1 SV=1	NSQEDSEDSKDVK
Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1 OS=Homo sapiens GN=NUCKS1 PE=1 SV=1	NSQEDSEDSKDVK
Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1 OS=Homo sapiens GN=NUCKS1 PE=1 SV=1	ATVTSPSPVK
Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1 OS=Homo sapiens GN=NUCKS1 PE=1 SV=1	EEDEEPESPPEK

Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1 OS=Homo sapiens GN=NUCKS1 PE=1 SV=1	EEDEEPESPPEKK
Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1 OS=Homo sapiens GN=NUCKS1 PE=1 SV=1	TSTSPPEK
Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1 OS=Homo sapiens GN=NUCKS1 PE=1 SV=1	KDDSHSAEDSEDEKEDHK
Nuclease-sensitive element-binding protein 1 OS=Homo sapiens GN=YBX1 PE=1 SV=3	NEGSESAPEGQAQQR
Nucleolar and coiled-body phosphoprotein 1 OS=Homo sapiens GN=NOLC1 PE=1 SV=2	VADNSFDAK
Nucleolar and coiled-body phosphoprotein 1 OS=Homo sapiens GN=NOLC1 PE=1 SV=2	ASSPFR
Nucleolar protein 56 OS=Homo sapiens GN=NOP56 PE=1 SV=4	EELMSSDLEETAGSTSIPK
Nucleolin OS=Homo sapiens GN=NCL PE=1 SV=3	KVVVSPTK
Nucleolin OS=Homo sapiens GN=NCL PE=1 SV=3	KVVVSPTKK
Nucleolin OS=Homo sapiens GN=NCL PE=1 SV=3	VVVSPTK
Nucleolin OS=Homo sapiens GN=NCL PE=1 SV=3	VVVSPTKK
Nucleolin OS=Homo sapiens GN=NCL PE=1 SV=3	KVVVSPTKK
Nucleophosmin OS=Homo sapiens GN=NPM1 PE=1 SV=2	SIRDTPAK
Nucleophosmin OS=Homo sapiens GN=NPM1 PE=1 SV=2	DELHIVEAEAMNYEGSPIK
Palladin OS=Homo sapiens GN=PALLD PE=1 SV=3	IASDEEIQGTK
PC4 and SFRS1-interacting protein OS=Homo sapiens GN=PSIP1 PE=1 SV=1	QSNASSDVEVEEK
Pinin OS=Homo sapiens GN=PNN PE=1 SV=4	EIAIVHSDAEK
Plasminogen activator inhibitor 1 RNA-binding protein OS=Homo sapiens GN=SERBP1 PE=1 SV=2	DELTESPK
Polycomb protein EED OS=Homo sapiens GN=EED PE=1 SV=2	CKYSFK
Prelamin-A/C OS=Homo sapiens GN=LMNA PE=1 SV=1	LRLSPSPTSQR
Prelamin-A/C OS=Homo sapiens GN=LMNA PE=1 SV=1	ASSHSSQTQGGGSVTK
Prelamin-A/C OS=Homo sapiens GN=LMNA PE=1 SV=1	LSPSPTSQR
Prelamin-A/C OS=Homo sapiens GN=LMNA PE=1 SV=1	LRLSPSPTSQR

Prelamin-A/C OS=Homo sapiens GN=LMNA PE=1 SV=1	LRLSPSPTSQR
Pre-mRNA-splicing factor 38B OS=Homo sapiens GN=PRPF38B PE=1 SV=1	RSLSPR
Probable 28S rRNA (cytosine(4447)-C(5))-methyltransferase OS=Homo sapiens GN=NOP2 PE=1 SV=2	GTDTQTPAVLSPSK
Probable 28S rRNA (cytosine(4447)-C(5))-methyltransferase OS=Homo sapiens GN=NOP2 PE=1 SV=2	GTDTQTPAVLSPSK
Prostaglandin E synthase 3 OS=Homo sapiens GN=PTGES3 PE=1 SV=1	DWEDDSDEDMSNFDR
Proteasome subunit alpha type-3 OS=Homo sapiens GN=PSMA3 PE=1 SV=2	ESLKEEDESDDDNM
Protein DEK OS=Homo sapiens GN=DEK PE=1 SV=1	EESEEEDEDEDEEEEEEEK
Protein ELYS OS=Homo sapiens GN=AHCTF1 PE=1 SV=3	SVENQESVEIINDLK
Protein ELYS OS=Homo sapiens GN=AHCTF1 PE=1 SV=3	EVSPSDVR
Protein ELYS OS=Homo sapiens GN=AHCTF1 PE=1 SV=3	EVSPSDVR
Protein FAM117B OS=Homo sapiens GN=FAM117B PE=1 SV=2	KKGSHK
Protein IWS1 homolog OS=Homo sapiens GN=IWS1 PE=1 SV=2	NQASDSENEELPKPR
Protein IWS1 homolog OS=Homo sapiens GN=IWS1 PE=1 SV=2	AAVLSDSEDEEK
Protein LSM14 homolog A OS=Homo sapiens GN=LSM14A PE=1 SV=3	SSPQLDPLR
Protein PAXX OS=Homo sapiens GN=C9orf142 PE=1 SV=2	LAAAEEETAVSPR
Protein SON OS=Homo sapiens GN=SON PE=1 SV=4	ESDQTLAALLSPK
Putative 40S ribosomal protein S10-like OS=Homo sapiens GN=RPS10P5 PE=5 SV=1	AEAGAGSATEFQFR
Ras GTPase-activating protein-binding protein 1 OS=Homo sapiens GN=G3BP1 PE=1 SV=1	SSSPAPADIAQTVQEDLR
Regulation of nuclear pre-mRNA domain-containing protein 2 OS=Homo sapiens GN=RPRD2 PE=1 SV=1	DVEDMELSDVEDDGSK
Regulator of chromosome condensation OS=Homo sapiens GN=RCC1 PE=1 SV=1	SPPADAIPK
Ribosomal L1 domain-containing protein 1 OS=Homo sapiens GN=RSL1D1 PE=1 SV=3	ATNESEDEIPQLVPIGK
Ribosomal L1 domain-containing protein 1 OS=Homo sapiens GN=RSL1D1 PE=1 SV=3	AAESETPGKSPEK
RNA-binding protein 10 OS=Homo sapiens GN=RBM10 PE=1 SV=3	LASDDRPSPPR
RNA-binding protein 39 OS=Homo sapiens GN=RBM39 PE=1 SV=2	IGLPHSIK

RNA-binding protein Raly OS=Homo sapiens GN=RALY PE=1 SV=1	GRLSPVPVPR
RNA-binding protein with serine-rich domain 1 OS=Homo sapiens GN=RNPS1 PE=1 SV=1	RFSPPR
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	KETSEAEADNLDDEK
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	KVELSESEEDK
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	AASPSPQSVR
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	KVELSESEEDKGGK
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	SPSPAPPPR
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	TASPPPPPK
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	HRPSPPATPPPK
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	SRVSVSPGR
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	RLSPSASPPR
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	RYSPPPPK
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	RVSHSPPPK
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	KETSEAEADNLDDEK
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	KVELSESEEDKGGK
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	EKTPELPEPSVK
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	AASPSPQSVR
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	APQTSSSPPVPR
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	KVELSESEEDK
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	TASPPPPPK
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	VSVSPGR
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	KEKTPELPEPSVK
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	TRHSPTQQSNR
Serine/arginine repetitive matrix protein 1 OS=Homo sapiens GN=SRRM1 PE=1 SV=2	KEKTPELPEPSVK
Serine/arginine repetitive matrix protein 2 OS=Homo sapiens GN=SRRM2 PE=1 SV=2	SGSSQELDVKPSASPQER

Serine/arginine repetitive matrix protein 2 OS=Homo sapiens GN=SRRM2 PE=1 SV=2	GRSPSPKPR
Serine/arginine repetitive matrix protein 2 OS=Homo sapiens GN=SRRM2 PE=1 SV=2	AGMSSNQSISSPVLDVPR
Serine/arginine repetitive matrix protein 2 OS=Homo sapiens GN=SRRM2 PE=1 SV=2	AQTTPGPSLSGSK
Serine/arginine repetitive matrix protein 2 OS=Homo sapiens GN=SRRM2 PE=1 SV=2	TSPPLDR
Serine/arginine repetitive matrix protein 2 OS=Homo sapiens GN=SRRM2 PE=1 SV=2	ENSFGSPLEFR
Serine/arginine repetitive matrix protein 2 OS=Homo sapiens GN=SRRM2 PE=1 SV=2	SRASPATHR
Serine/arginine repetitive matrix protein 2 OS=Homo sapiens GN=SRRM2 PE=1 SV=2	ENSFGSPLEFR
Serine/arginine repetitive matrix protein 2 OS=Homo sapiens GN=SRRM2 PE=1 SV=2	ELNSPLR
Serine/arginine repetitive matrix protein 2 OS=Homo sapiens GN=SRRM2 PE=1 SV=2	SRLSLR
Serine/arginine-rich splicing factor 11 OS=Homo sapiens GN=SRSF11 PE=1 SV=1	DYDEEEQGYDSEKEK
Serine/arginine-rich splicing factor 11 OS=Homo sapiens GN=SRSF11 PE=1 SV=1	KPIETGSPK
Serum response factor-binding protein 1 OS=Homo sapiens GN=SRFBP1 PE=1 SV=1	SLDFPQNEPQIK
Serum response factor-binding protein 1 OS=Homo sapiens GN=SRFBP1 PE=1 SV=1	EYFDDSTEER
Signal-induced proliferation-associated 1-like protein 1 OS=Homo sapiens GN=SIPA1L1 PE=1 SV=4	TLSDESIYNSQR
Signal-induced proliferation-associated 1-like protein 1 OS=Homo sapiens GN=SIPA1L1 PE=1 SV=4	LIDLESPTPESQK
Small acidic protein OS=Homo sapiens GN=SMAP PE=1 SV=1	SASPDDDLGSNWEAADLGNEER
Sororin OS=Homo sapiens GN=CDCA5 PE=1 SV=1	APSPTKPLR
Spectrin beta chain, non-erythrocytic 1 OS=Homo sapiens GN=SPTBN1 PE=1 SV=2	RPPSPEPSTK
Splicing factor 45 OS=Homo sapiens GN=RBM17 PE=1 SV=1	RDPDSDDEDYER
Src substrate cortactin OS=Homo sapiens GN=CTTN PE=1 SV=2	TQTPPVSPAPQPTEER
Src substrate cortactin OS=Homo sapiens GN=CTTN PE=1 SV=2	LPSSPVYEDAASF
Stathmin OS=Homo sapiens GN=STMN1 PE=1 SV=3	ESVPEFPLSPPK
Suppressor of SWI4 1 homolog OS=Homo sapiens GN=PPAN PE=1 SV=1	VGGSDEEASGIPSR
Target of EGR1 protein 1 OS=Homo sapiens GN=TOE1 PE=1 SV=1	AADSDDGAVSAPAASDGGVSK

Targeting protein for Xklp2 OS=Homo sapiens GN=TPX2 PE=1 SV=2	SSDQPLTVPVSPK
THO complex subunit 4 OS=Homo sapiens GN=ALYREF PE=1 SV=3	QQLSAEELDAQLDAYNAR
Thyroid hormone receptor-associated protein 3 OS=Homo sapiens GN=THRAP3 PE=1 SV=2	IDISPSTFR
Thyroid hormone receptor-associated protein 3 OS=Homo sapiens GN=THRAP3 PE=1 SV=2	ASAVSELSPR
Transcription factor AP-1 OS=Homo sapiens GN=JUN PE=1 SV=2	AKNSDLLTSPDVGLLK
Transcription factor AP-1 OS=Homo sapiens GN=JUN PE=1 SV=2	LASPELER
Transcription factor jun-B OS=Homo sapiens GN=JUNB PE=1 SV=1	DATPPVSPINMEDQER
Transcription initiation factor TFIID subunit 12 OS=Homo sapiens GN=TAF12 PE=1 SV=1	LSPENNQVLTK
Transcriptional repressor protein YY1 OS=Homo sapiens GN=YY1 PE=1 SV=2	EEVVGGDDSDGLR
Treacle protein OS=Homo sapiens GN=TCOF1 PE=1 SV=3	TSQVGAASAPAKESPR
Treacle protein OS=Homo sapiens GN=TCOF1 PE=1 SV=3	AALAPAKESPR
Treacle protein OS=Homo sapiens GN=TCOF1 PE=1 SV=3	LGAGEGGEASVSPEK
Treacle protein OS=Homo sapiens GN=TCOF1 PE=1 SV=3	KLSGDQPAAR
U4/U6.U5 tri-snRNP-associated protein 1 OS=Homo sapiens GN=SART1 PE=1 SV=1	RVSEVEEEKEPVPQPLPSDDTR
U4/U6.U5 tri-snRNP-associated protein 2 OS=Homo sapiens GN=USP39 PE=1 SV=2	EVDEDESEPER
UAP56-interacting factor OS=Homo sapiens GN=FYTDD1 PE=1 SV=3	LVGATATSSPPPK
Uncharacterized protein C15orf52 OS=Homo sapiens GN=C15orf52 PE=1 SV=1	SPPTQVAISSDSAR
Uncharacterized protein CXorf67 OS=Homo sapiens GN=CXorf67 PE=2 SV=1	SGSPDPEVPSR
Zinc finger CCCH domain-containing protein 14 OS=Homo sapiens GN=ZC3H14 PE=1 SV=1	DLVQPDKPASPK
Zinc finger protein 185 OS=Homo sapiens GN=ZNF185 PE=1 SV=3	QSSPSGSEQLVR
Zinc finger protein 22 OS=Homo sapiens GN=ZNF22 PE=1 SV=3	FDSSFSR
Zinc finger protein 579 OS=Homo sapiens GN=ZNF579 PE=1 SV=2	ESESEAEAGAAELR
Zinc finger protein 740 OS=Homo sapiens GN=ZNF740 PE=1 SV=1	AGSPDVLR

Appendix II: Hatimy, A.A. et al., 2015. Histone H2AX Y142 phosphorylation is a low abundance modification. *International Journal of Mass Spectrometry*, 391, pp.1–7.

International Journal of Mass Spectrometry 391 (2015) 139–145



Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

journal homepage: www.elsevier.com/locate/ijms



Histone H2AX Y142 phosphorylation is a low abundance modification



Abubakar A. Hatimy^a, Martin J.G. Browne^b, Andrew Flaus^b, Steve M.M. Sweet^{a,*}

^a Genome Damage and Stability Centre, University of Sussex, Brighton, UK

^b Centre for Chromosome Biology, School of Natural Sciences, National University of Ireland, Galway, Ireland

ARTICLE INFO

Article history:

Received 22 May 2015

Received in revised form 20 July 2015

Accepted 28 July 2015

Available online 5 August 2015

Keywords:

Chromatin

Histone

Phosphorylation

DNA damage

Selected reaction monitoring

ABSTRACT

We employ targeted mass spectrometry to compare the levels of H2AX S139 phosphorylation (γ H2AX) and Y142 phosphorylation. We use synthetic peptides to facilitate MS optimisation and estimate relative detection efficiencies for the different modifications. Despite phosphopeptide enrichment from large amounts of starting material, we are unable to detect endogenous H2AX Y142 phosphorylation, indicating that it is present in low abundance ($<1\%$). We also calculate the relative levels of H2AX compared to other H2A isoforms and quantify the proportion of H2AX that is phosphorylated on S139 (γ H2AX) after ionising radiation.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

1.1. S139 phosphorylation

The histone H2A family variant, H2AX, is distinguished from canonical H2A family members through a 22 amino acid C-terminal tail [1]. Phosphorylation of the C-terminal domain of H2AX at position 139 (γ H2AX) is a rapid response to DNA double-strand breaks (DSB). S139 is phosphorylated by ATM, ATR and DNA-PK, which are phosphatidylinositol 3-kinase-related kinases. γ H2AX foci are widely used as diagnostic markers of DSB. The utility of γ H2AX as a marker stems from the rapid (<1 min) and extensive nature of this modification. Rogakou et al. observed that approximately 1% of total H2AX becomes phosphorylated per gray of ionising radiation (IR), and extrapolated from H2AX relative abundance that each DSB results in γ H2AX covering on average 2 million bp [2]. The biological function of such large γ H2AX domains is not clear, and the H2AX histone is not essential for DSB repair; however, H2AX^{-/-} mice show increased ionising radiation sensitivity, as well as increases in chromatid breaks and dicentric chromosomes [3].

1.2. Y142 phosphorylation

The H2AX C-terminal domain can also be phosphorylated on tyrosine 142 by the WSTF remodelling factor kinase [4–6]. Cook

et al. show that dephosphorylation of Y142 upon DNA damage avoids apoptosis. Using synthetic phosphopeptides, they demonstrate binding of pro-apoptotic factors to S139 Y142 doubly phosphorylated peptides: the implication is that Y142 phosphorylation is abundant, and will be located in proximity to DNA damage. While kinases and phosphatases responsible for creating and removing this modification have been identified, the basal level of Y142 phosphorylation is unknown, although our earlier intact histone MS analysis indicates that in HeLa cells it is not greater than $\sim 10\%$ [7]. Scully et al. expressed epitope-tagged H2AX in H2AX^{-/-} mouse ES cells and identified a number of H2AX modifications by mass spectrometry, including S139 and T101 phosphorylation; however, Y142 phosphorylation was not detected [8]. The role of Y142ph in the DNA damage response is of great interest, with the identification of putative interacting proteins that recognise the doubly phosphorylated C-terminal tail [9]. Mutation of these residues has been carried out in the chicken DT40 cell line and revealed that Y142A IR sensitivity is rescued by co-mutation of S139A [10].

1.3. H2AX levels across cell lines and in the genome

Given the role of H2AX phosphorylation in signalling DNA damage and recruiting repair proteins, it is interesting to consider the levels of H2AX present across the genome in human cells. Rogakou et al. employed two-dimensional gel separation of histone preparations with Coomassie staining to quantify H2AX as a proportion of all H2A1, H2A2 and H2AX signal. They analysed four human cell lines and found that H2AX makes up 2–25% of the total [2].

* Corresponding author. Tel.: +44 1273877414.
E-mail address: s.m.sweet@sussex.ac.uk (S.M.M. Sweet).

2. Materials and methods

2.1. Synthetic peptides and recombinant histones

Peptides listed below were synthesised by JPT (Germany).

Synthetic peptide sequence		
GHYAERATQASQEFY	KGHYAERGTGGKAR	GKTGGKARAKAKSR
GHYAERATQA-PS-QEY	KGNYAERGTGGKAR	GKQGGKARAKAKSR
GHYAERATQASQEF-pY-		

Recombinant H2AX and H2A1B were expressed in *E. coli* Rosetta 2(DE3) pLysS. Cultures were incubated at 37 °C with 180 rpm agitation. Unlabelled proteins were grown in 2YT media to OD₆₀₀ 0.6 and induced with isopropyl-β-D-thiogalactoside (IPTG) for 4 h before harvest. ¹⁵N-labelled proteins were produced using a two-step expression procedure. 2YT cultures were grown to OD₆₀₀ 0.6 and cells collected by centrifugation. Cells were gently resuspended in an equivalent volume of pre-warmed minimal media (50 mM Na₂HPO₄, 50 mM KH₂PO₄, 1 g/L (¹⁵NH₄)₂SO₄, 2 mM MgSO₄, 20 mM citric acid, 30 mg/L thiamine, pH 7.0) and incubated at 37 °C for 30 min prior to IPTG induction for 4 h and harvest. Recombinant histones were purified according to the established inclusion body wash procedure followed by cation exchange chromatography [11]. Purified histones were dialysed into water and lyophilised for long-term storage. The recombinant proteins were resuspended in water at a concentration of 2 mg/ml, based on absorbance at 278 nm using extinction coefficients 5960 and 4470 M⁻¹ cm⁻¹ for H2AX and H2A1B, respectively. MS analysis of ¹⁵N H2AX indicated that labelling was >98% complete. Recombinant histone sequences are shown below. Synthetic peptide sequences are underlined.

H2AX: MSGRGKTKGGKARAKAKSRSSRAGLQFPVGRVHRLLRKGH-YAERVGAGAPVYLAADVLEYLTAEILELAGNAARDNKKTRIIPRHLQLAIR-NDEELNKLGGVTIAQGGVLPNIQAVLLPKKTSATVGPAPSGGKKAT-QASQEFY

H2A1B: MSGRGKTKGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGNYSERVGAGAPVYLAADVLEYLTAEILELAGNAARDNKKTRIIPRHLQLAIR-NDEELNKLGRVTIAQGGVLPNIQAVLLPKKTESHHKAKGK

Cell lines and cell culture: Human U2-OS and HeLa cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% foetal calf serum, 2 mM L-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin. Epstein Barr virus-transformed wild-type lymphoblastoid cells (LCLs; a gift from Mark O'Driscoll; [12]) were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 15% foetal calf serum, 2 mM L-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin. Approximately 6 × 10⁷ cells were grown for each sample.

Irradiating cells: U2-OS cells were irradiated with a dosage of 20 Gy using a ¹³⁷Cs source. Cells were immediately placed in a 37 °C incubator and were left to recover for either 5 min, 30 min, or 60 min. Upon recovery, cells were spun down and were immediately flash frozen and stored at -80 °C before sample preparation.

Nuclei isolation and acid extraction: Nuclei Isolation Buffer 250 (NIB-250) was prepared: 15 mM Tris HCl (pH 7.5), 60 mM KCl, 15 mM NaCl, 5 mM MgCl₂, 1 mM CaCl₂, and 250 mM Sucrose. NIB-250 was supplemented with the following phosphatase and protease inhibitors: 10 mM Na-butyrate, 0.5 mM AEBSEF, 2 µM leupeptin, 1 µM Pepstatin A, 1 µM aprotinin, 2 mM PMSF, 50 mM NaF, 2 mM Na-orthovanadate, and 60 mM β-glycerophosphate. NIB-250 with 0.3% NP-40 was added to the cells at a ratio of 10:1. Cells were incubated on ice for 5 min and then subsequently spun down at 600 rcf for 5 min at 4 °C. The supernatant was discarded and then NIB-250 without NP-40 was added to the cell pellet and spun down again, twice. The remaining pellet consisted of isolated nuclei. The

isolated nuclei were vortexed slowly and 0.4 N sulphuric acid was added to the cells at a final ratio of 5:1. The nuclei-H₂SO₄ suspension was left to incubate on ice for 60 min and then spun down at the centrifuge's maximum speed for 5 min at 4 °C. The supernatant was transferred to a microcentrifuge tube. To the supernatant, 100% trichloroacetic acid (TCA) was added to give a final concentration of 20% TCA. The mixture was left to precipitate overnight at 4 °C. The precipitate was then spun down at the microcentrifuge's maximum speed. The pellet was resuspended in chilled acetone with 0.1% hydrochloric acid. The sample was then spun down, washed twice with acetone, left to air dry, and finally dissolved in water.

Derivatisation and trypsinisation: A modified version of the derivatisation protocol described in Maile et al. was employed [13]. From the acid extracted histone sample, 5 µg was used to make up a volume of 9 µL. 1 µL of 1 M tetraethylammonium bromide (TEAB) was added to the sample to buffer the subsequent reaction. 1 µL of diluted propionic anhydride (in water to a ratio of 1:100) was added to the histones. The samples were vortexed for 2 min at room temperature. To quench the derivatisation reaction, 1 µL of 80 mM hydroxylamine was added and the samples were left to incubate for 20 min at room temperature. The pH of the samples was measured to ensure optimal trypsin digestion, ~pH 8.5. 250 ng of trypsin was added to each sample and left to digest at 37 °C for 4 h or overnight. Upon digestion, a second round of derivatisation was carried by adding 3 µL of 1:100 diluted propionic anhydride. The samples were vortexed for 2 min at room temperature. 3 µL of 80 mM hydroxylamine was then added to quench the second round of derivatisation.

Phosphopeptide enrichment: TiO₂ metal oxide affinity chromatography was approximately as described in Matheron et al. [14]. Briefly, TiO₂ beads (Titansphere; 5 µm; GL Sciences, Japan) were resuspended at 10 mg/ml in 30% acetonitrile (ACN), 0.1% TFA. 50 µL of this slurry was packed by centrifugation (400 × g) into a GELoader tip spin column (Eppendorf; 0.5–20 µL volume) fritted with a C8 filter membrane (Empore). The spin column was equilibrated with loading buffer (80% ACN, 6% TFA; 50 µL at 2200), the sample was loaded (2 × 50 µL; 200 × g) and the column washed twice (50 µL; 50% ACN, 0.5% TFA, 0.2 M NaCl; 500 × g), with salt omitted in the second wash. Phosphopeptides were eluted with 20 µL of 10% NH₄OH (200 × g), into 25 µL 10% FA. A final elution of 3 µL 80% ACN, 2% FA was employed to release peptides from the C8 frit. Further acid (3 µL FA) was added to the eluates. Eluates were dried in a vacuum centrifuge and resuspended in 10 µL 0.1% TFA.

Nano-LC/MS: Peptide samples were analysed by nano-LC-MS (ThermoFisher U3000 nanoLC and Orbitrap XL mass spectrometer). Peptides were loaded onto a C18 trapping cartridge (Pepmap100 C18; 0.3 × 5 mm i.d.; 5 µm particle size) for 5 min at a flow-rate of 5 µL/min in 0.1% TFA loading buffer. Peptides were separated on an analytical column (PepMap100; 25 cm × 75 µm; 5 µm particle size) by a gradient from 1 to 35% ACN over 28 min, in the presence of 0.1% FA, at a flow rate of 0.3 µL/min.

Time	0	15	25	28	28	35	35
% acetonitrile	1	12	35	60	90	90	1

Nanospray was from a New Objective emitter with 10 µm tip (FS360-20-10-N-20). Pseudo-SRM was carried out in the linear ion trap of an Orbitrap XL, with a precursor isolation window of 2 m/z, an ion-trap fill time of maximum 50 ms and an LTQ target ion count of 1E4 [15]. A high-resolution precursor scan was carried out in the Orbitrap (SE5 target). Total cycle time was <30 s, enabling sufficient points across eluting peptide peaks for quantitation. For the TiO₂-enrichment eluates, in addition to the targeted LTQ events, a data-dependent event was included, to enable identification of untargeted/enriched phosphopeptides.

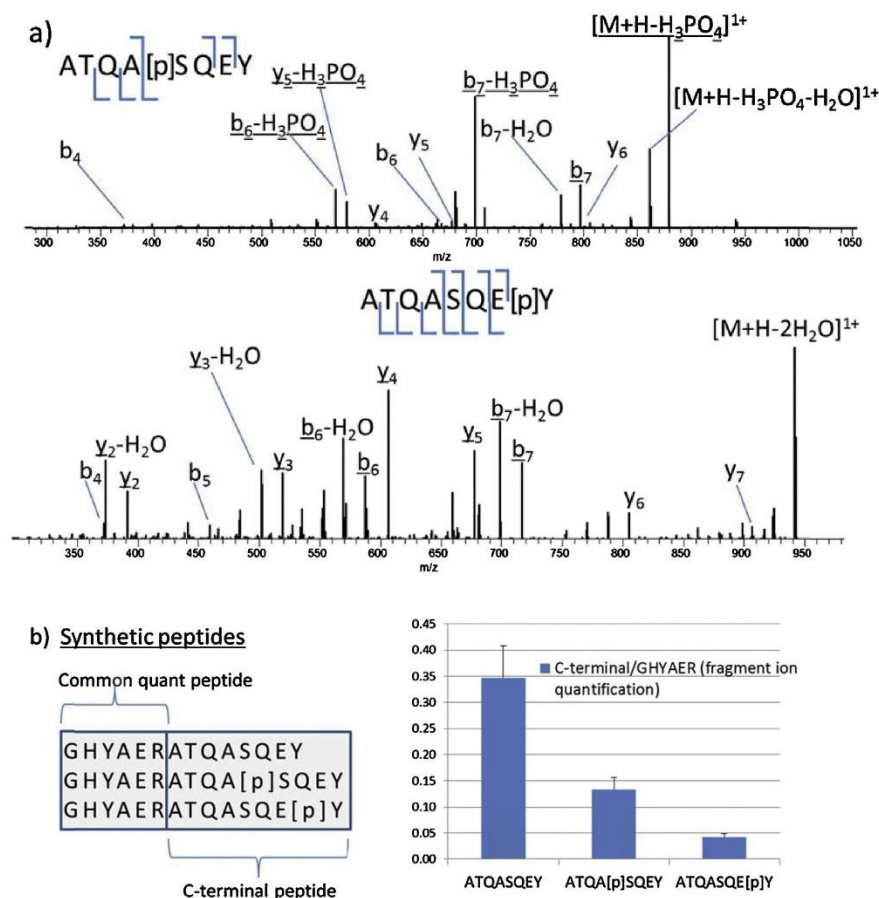


Fig. 1. (a) MS/MS fragmentation of singly charged ATQASQEY S139 or Y142 phosphorylated precursors. Fragment ions used for quantitation are underlined. (b) Concatenated synthetic peptides allow assessment of relative detection efficiency after trypsin digestion, nanoLC separation and pseudo-SRM peptide fragmentation. Both N-terminal and C-terminal peptides are unique to H2AX. Error bars show standard deviation ($n=3$).

Pseudo-SRM creation and analysis: Skyline v3.1 (MacCoss Lab, University of Washington) was used for both development of pseudo-SRM methods and for data analysis [16]. Peptide sequences for protein of interest were obtained from Uniprot and entered into Skyline. Putative and known post-translational modifications of interest were added to the peptides. Predicted b and y ions were surveyed on the instrument. To distinguish between different isobaric masses at least three transitions were selected that were unique to each peptide. Transitions used for data analysis are listed in Supplementary Table 1.

2.2. Data analysis

Database searching for the TiO₂-enriched samples employed Mascot and the Swissprot database. Settings were as follows: two missed cleavages; precursor 5 ppm error; fragment 0.6 Da error; variable modifications of acetylation (K), mono-, di- and trimethylation (K), phosphorylation with neutral loss (S,T), phosphorylation (Y). Three U2-OS acid extract TiO₂ eluates (25 µg, 50 µg, 50 µg + Y142ph peptide) were run as a combined search. Mascot results were filtered to require expectation value <0.05 and

bold red and exported as .csv. The peptide list was sorted according to descending *E*-value and duplicate pep_scan_titles were removed. 1393 peptide identifications remained. Duplicates in all three of pep_seq and pep_var_mod and pep_var_mod_pos were removed (again removing lowest *E*-value duplicates). 456 identifications remained. 307 of these were phosphorylated (including spiked Y142ph peptide). Some were clearly false-positives (as expected given *E*-value 0.05): filtering to require *E*-value <0.01 leaves 370 identifications, of which 230 were phosphorylated.

3. Results

In order to develop sensitive selected reaction monitoring (SRM) assays for H2AX phosphopeptide detection by mass spectrometry, as employed previously for histone H3 methylation analysis [17], we designed synthetic peptides containing either phosphorylated Y142, phosphorylated S139 or no modification. To allow determination of relative ionisation efficiency, these peptides were synthesised concatenated with a common quantification peptide. After trypsin digestion to separate the two peptides, the released phosphopeptides were fragmented (Fig. 1a), and the indicated

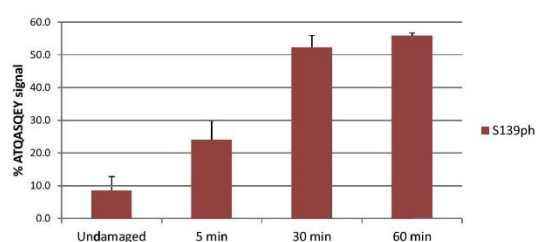


Fig. 2. Quantification of γ H2AX Ser139 phosphorylation after DNA damage. (A) Time-course after 20 Gy IR (U2-OS). Fragment ions from the singly charged unmodified and S139 phosphorylated precursors were quantified. Mean of two biological replicates is shown, each of which was analysed in technical triplicate. Error bars show standard deviation. Data shown are uncorrected for detection efficiency.

fragments were quantified in pseudo-SRM assays utilising the linear ion trap of an Orbitrap XL [15]. The H2AX C-terminal peptides were detected in both singly and doubly charged forms. The singly charged species was preferred for sensitive detection of low levels of C-terminal peptide in the presence of abundant canonical histone peptides, due to low background in this m/z region for these transitions. The released H2AX-specific quantification peptide (GHYAER) acts as an internal standard: the ionisation efficiency of the C-terminal peptide is measured relative to the quantification peptide in each case (Fig. 1b). We found that the phosphorylated peptides were detected $2.6\times$ and $8.3\times$ less well than the unphosphorylated peptide (S139ph and Y142ph, respectively; Fig. 1b).

We apply this correction factor to data obtained from U2-OS cells after ionising radiation (IR) to estimate that approximately 75% of H2AX in the sample was phosphorylated at S139 60 min after 20 Gy IR (Fig. 2).

3.1. Y142 phosphorylation

We did not detect Y142ph in undamaged cells or any of the time-points shown in Fig. 2. We considered that the failure to detect Y142 phosphorylation in our histone preparations could indicate that this modification is present only at low abundance, given that we were able to detect the synthetic Y142ph peptide with reasonable sensitivity. We therefore carried out a TiO_2 metal oxide affinity phosphopeptide enrichment step on a histone preparation from undamaged U2-OS cells, which we had shown to contain only low levels of γ H2AX phosphorylation (approximately 4%). We carried out a parallel enrichment on a control sample, containing 50 pmol BSA and 0.5 pmol of α -casein, Y142ph, S139ph and unmodified synthetic peptides (Fig. 3). We see efficient enrichment of Y142ph, S139ph from the control sample, in addition to the expected α -casein phosphopeptides. We also see efficient enrichment of S139ph from the undamaged U2-OS sample, alongside a number of other phosphopeptides from other histones and low-level phosphoproteins present in the acid-extract. However, no Y142ph signal was detected from the U2-OS sample.

To further verify that there was no technical problem with our enrichment, we spiked Y142ph peptide (2 pmol) into a U2-OS sample (50 μg acid extract). Enrichment from the spiked sample was carried out in parallel with enrichment from an equivalent unspiked sample. As before, endogenous S139ph was enriched in both cases. Y142ph was successfully enriched from the spiked sample, but not detectable in the unspiked sample (Fig. 4). Although the U2-OS acid extracts contain mainly histones, TiO_2 enrichment allowed identification of 232 distinct phosphopeptides (with an expectation value <0.01), mainly from low abundance non-histone proteins (Supp. Table 2). These identifications included the previously described CDK3 Y15 tyrosine phosphorylation. In addition to H2AX S139ph, previously reported core histone phosphorylations H3 S10ph, S57ph and H4 S47ph were also identified, as well as nine H1 sites.

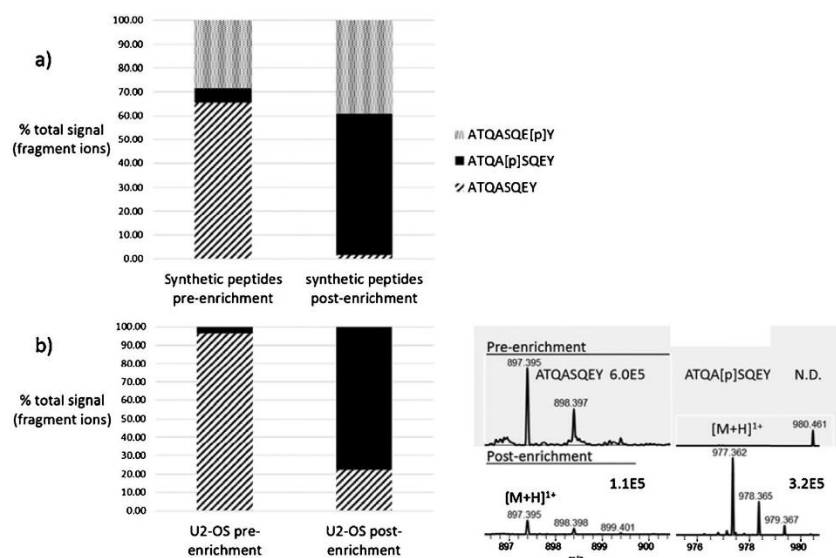


Fig. 3. TiO_2 -based phosphopeptide enrichment of H2AX C-terminal peptides. (A) Successful enrichment of S139ph and Y142ph peptides from 100 \times excess BSA peptides. (B) Enrichment of endogenous S139ph peptide from undamaged U2-OS, relative to the unmodified peptide. LHS: quantification of fragment ions from singly charged precursors. RHS: precursor ions (with relative intensity shown).

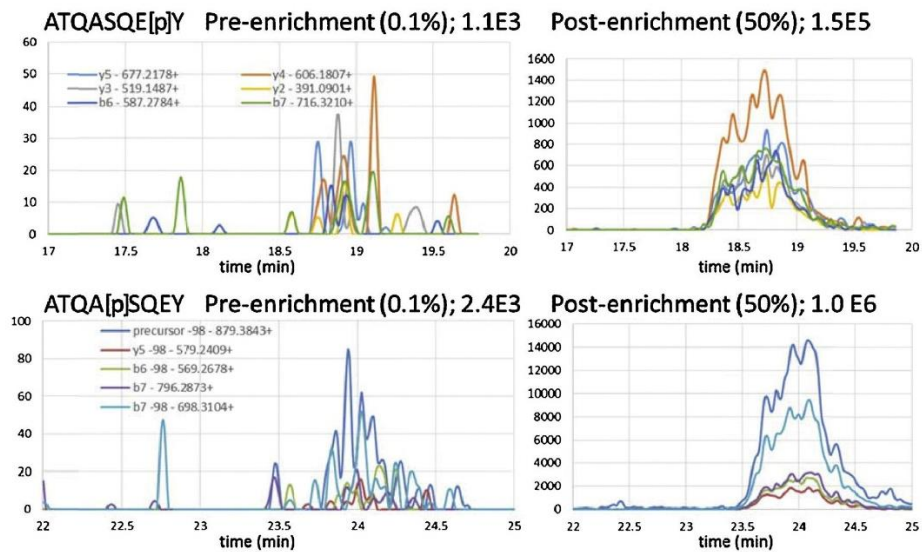


Fig. 4. Successful enrichment of Y142ph synthetic peptide spiked into U2-OS acid extract. 0.1% of the 50 μ g sample was analysed pre-enrichment; both Y142ph (synthetic) and S139ph are barely detectable using pseudo-SRM (precursor ions are not detectable; data not shown). After enrichment both H2AX phosphopeptides are detected with good signal-to-noise.

3.2. H2AX abundance across cell lines

We compared the abundance of two peptides unique to H2AX (GKTGGKAR and KGHYAER) to similar peptides in other H2A family members in three cell lines: U2-OS, HeLa and an EBV-transformed lymphoblastoid cell line (LCL; Table 1). In this analysis we employed a propionylation derivatisation step prior to, and after, trypsin digestion. In order to control for possible variations between sample preparations, 15 N-labelled recombinant H2AX and H2A1B were spiked into the three samples. These samples were then derivatized

and digested as normal and analysed in technical triplicate. Fig. 5 shows the normalised ratio of endogenous H2AX to 15 N H2AX peptide for each cell line, indicating that U2-OS contains the highest relative abundance of H2AX, $>2\times$ more than HeLa. The two H2AX peptides both give similar results.

3.3. H2AX abundance compared to other H2A family variants

We quantified the signal from the two peptides unique to H2AX as a proportion of the signal from closely related H2A family

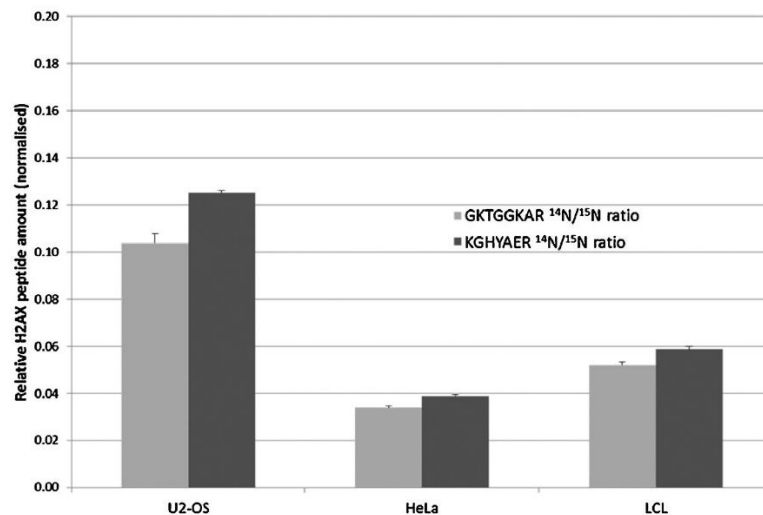


Fig. 5. Relative amount of H2AX across three cell lines. 15 N recombinant H2AX and H2A1B were spiked into U2-OS, HeLa and LCL extracts; ratios for diagnostic peptides were calculated and normalised to ratio of H2A1B GKQGGKAR, to control for small differences in starting acid extract amounts. Error bars show standard deviation.

Table 1

H2AX abundance relative to H2A family members. GKTGGKAR was found to be detected 1.1× better than the canonical version, GKQGGKAR. The H2AX peptide KGHYAER was detected 1.7× better than the canonical version KGNYAER. The proportions are corrected accordingly in the table. Data are from two biological replicates, analysed in technical triplicate. The standard deviation is shown in brackets.

Measured peptide	Protein(s)	U2-OS	HeLa	LCL
GKTGGKAR	H2AX	4 (0.7)	1 (0.6)	2 (0.03)
GKQGGKAR	11 H2A (1A, 1B, 1C, 1H, 1J, 1D, 1, 2A, 2C, 3, 2B)	96 (0.6)	98 (0.9)	98 (0.2)
GKQGGKVR	H2A]	1 (0.05)	1 (0.3)	0 (0.1)
Not measured	H2AB1, H2AB2, H2AV, H2AZ	n.m.	n.m.	n.m.
Total	17	100	100	100
KGHYAER	H2AX	6 (0.8)	3 (0.3)	5 (0.5)
KGNYAER	9 H2A (1A, 1C, 1H, 1J, 1, 2A, 2C, 2B, J)	87 (4.2)	88 (3.1)	83 (5.7)
KGNYSER	H2A 1B, 1D, 3	7 (5.0)	9 (2.8)	12 (6.2)
Not measured	H2AB1, H2AB2, H2AV, H2AZ	n.m.	n.m.	n.m.
Total	17	100	100	100
	H2AX average	5	2	4

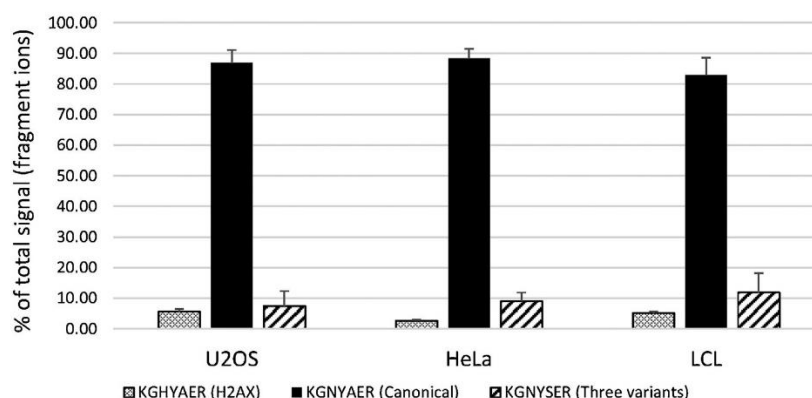


Fig. 6. Quantification of H2AX abundance relative to other H2A isoforms. Relative H2AX levels (corrected for detection efficiency) from U2-OS, HeLa and LCL. Quantification is of fragment ions from doubly charged precursors. Data are from two biological replicates, analysed in technical triplicate. KGHYAER is unique to H2AX; and KGNYSER is unique to H2A 1B, 1D and 3. Error bars show standard deviation.

peptides (Fig. 6 and Table 1). Concatenated peptides were again employed to correct for detection efficiencies, as above. The same trend with respect to H2AX abundance across the three cell lines is observed, as expected (U2-OS > LCL > HeLa). The estimates of H2AX abundance are not identical for the two different H2AX peptides. The proportion of H2AX calculated from the KGHYAER peptide is consistently higher than that calculated from the GKTGGKAR peptide. This difference may be due to the different ionisation properties of the non-canonical peptides (KGNYSER and GKQGGKVR), synthetic versions of which were not available for correction. Importantly, both peptides give a highly similar relative abundance across the three cell lines. The values shown in Table 1 represent upper estimates, as four divergent H2A variants are not measured at all.

4. Discussion

We have developed pseudo-SRM assays to quantify levels of H2AX C-terminal phosphorylation. The C-terminal phosphopeptides of H2AX are not the most straightforward peptides to analyse in positive-mode mass spectrometry, being without basic residues to balance the negative charge of the phosphate group. Nonetheless, both S139ph and Y142ph peptides ionise to give doubly and singly charged species, which can be fragmented to give diagnostic fragment ions. The two phosphopeptides show considerable difference in retention time, avoiding co-fragmentation of the

isobaric species. We detect variable levels of γ H2AX (S139ph) from undamaged U2-OS (4% and 12% in the two biological replicates contributing to Fig. 2), increasing to 75% 1 h after 20 Gy IR and correction for detection efficiency. We note that this level of S139ph is substantially higher than that previously reported by Rogakou et al. (30–50%), albeit using a different method and for different cell lines [2]. Our correction factor controls for ionisation efficiency and fragmentation efficiency of the unmodified and phosphorylated peptides; however, it does not control for differences in trypsin release of the measured peptides. We consider it possible that phosphorylation of S139 may affect the ability of trypsin to cleave the proximal lysine, producing the quantified fully tryptic peptide: a set of synthetic peptides incorporating the endogenous trypsin cleavage site would enable this to be tested (APSGGKKATQASQEQY). The 75% estimate should therefore be treated with caution.

We were unable to detect endogenous H2AX Y142ph, even after enrichment of phosphopeptides from large amounts of histone preparations (50 μ g). Notably, synthetic Y142ph was detected and enriched without difficulty. Given that after enrichment we identify endogenous S139ph from undamaged cells (approximately 5% abundance) with a signal to noise ratio of over 100, we should be able to identify S139ph at 20–40× lower abundance, i.e. 0.25–0.125% abundance within the genome. Our synthetic peptide data indicate that the Y142ph peptide is detected 3.2× less well than the S139ph peptide, implying a conservative detection level of 0.8% overall abundance. Large-scale phosphoproteomic studies

have similarly failed to detect Y142ph, even when S139 phosphorylation was detected, and in the absence of DNA damage treatments [18,19]. We conclude that the stoichiometry of Y142ph is low, at least below 1% of H2AX in undamaged U2-OS. Of the early papers describing H2AX Y142ph, Krishnan et al. also used U2-OS cells [6]. It is possible that Y142ph levels are higher in the other cell lines studied previously (HEK 293T, mouse 3T3 and mouse embryonic fibroblasts). However, if Y142 phosphorylation plays a significant physiological role in apoptosis and the DNA damage response, it seems unlikely to vary dramatically across cell lines or between human and mouse.

This low stoichiometry has implications for the use of Y142ph antibodies: Xiao et al. raised an antibody that appears to strongly prefer Y142ph, but nevertheless still binds unphosphorylated Y142 (ELISA Xiao et al. Supp. Fig. 1). With >100:1 ratio of unmodified to Y142ph, this background could be problematic [4]. The low abundance of Y142ph also has implications for models involving binding of proteins to both S139ph and Y142ph simultaneously [5,9]: co-localisation of S139ph and Y142ph on the same H2AX C-terminal tail would be most likely to occur at extremely high levels of DNA damage, when the majority of H2AX S139 are phosphorylated. Analysis of the role of Y142 through site-directed mutagenesis [10] is complicated by the observation that Y142 mutation interferes with binding of proteins such as MDC1 to γ H2AX [20].

Removal of WSTF, the kinase responsible for Y142 phosphorylation, reduces γ H2AX formation and abrogates the DNA damage response; however, it is unclear if this is related to H2AX Y142 phosphorylation or to the chromatin remodelling role of this factor [4]. Our results would be consistent with an indirect effect.

We analyse H2AX abundance in three cell lines, and find that H2AX constitutes between 2 and 5% of the thirteen H2A family members quantified: the total level would therefore be lower when the remaining four, more divergent, variants are taken into account (H2AV, H2AZ, H2AB1, H2AB2). This level seems reasonable, considering that the gene encoding H2AX, H2AFX, is one of 24 H2A family (H2AF) genes, encoding 17 unique H2A isoforms (Uniprot curated protein family). Therefore, if all expression levels were even, H2AX would make up 4.2% of cellular H2A.

Our estimate for HeLa of 1.8% H2AX is very similar to that of Rogakou et al. [2] (2.4%). Presumably the divergent H2A family members do not co-migrate with the canonical H2As upon two-dimensional gel electrophoresis, and therefore, were also not quantified by Rogakou et al.

Acknowledgements and funding

The authors thank Dr Rita Colnaghi and Professor Mark O'Driscoll for advice and the generous gift of the LCL cells; Dr Lisa Woodbine for advice on cell culture; Dr Stuart Rulten for advice and assistance with IR treatment; and Dr Shabaz Mohammed (University of Oxford) for advice on phosphopeptide enrichment. AAH is funded by a School of Life Sciences PhD studentship and SMMS is funded by an MRC career development award.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ijms.2015.07.028>

References

- [1] D. Pinto, A. Flaus, Structure and Function of Histone H2AX, in: H.-P. Nasheuer (Ed.), *Genome Stability and Human Diseases*, Springer, Netherlands, 2010, pp. 55–78.
- [2] E.P. Rogakou, D.R. Pilch, A.H. Orr, V.S. Ivanova, et al., DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139, *J. Biol. Chem.* 273 (1998) 5858–5868.
- [3] A. Celeste, S. Petersen, P.J. Romanienko, O. Fernandez-Capetillo, et al., Genomic instability in mice lacking histone H2AX, *Science* 296 (2002) 922–927.
- [4] A. Xiao, H. Li, D. Shechter, S.H. Ahn, et al., WSTF regulates the H2AX DNA damage response via a novel tyrosine kinase activity, *Nature* 457 (2009) 57–62.
- [5] P.J. Cook, B.G. Ju, F. Telesse, X. Wang, et al., Tyrosine dephosphorylation of H2AX modulates apoptosis and survival decisions, *Nature* 458 (2009) 591–596.
- [6] N. Krishnan, D.G. Jeong, S.-K. Jung, S.E. Ryu, et al., Dephosphorylation of the C-terminal tyrosyl residue of the DNA damage-related histone H2AX is mediated by the protein phosphatase ppp1, *J. Biol. Chem.* 284 (2009) 16066–16070.
- [7] J.C. Tran, L. Zamborg, D.R. Ahlf, J.E. Lee, et al., Mapping intact protein isoforms in discovery mode using top-down proteomics, *Nature* 480 (2011) 254–258.
- [8] A. Xie, S. Odate, G. Chandramouly, R.A. Scully, H2AX post-translational modifications in the ionizing radiation response and homologous recombination, *Cell Cycle* 9 (2010) 3602–3610.
- [9] N. Singh, H. Basnet, T.D. Wiltshire, D.H. Mohammad, et al., Dual recognition of phosphoserine and phosphotyrosine in histone variant H2AX by DNA damage response protein MCPH1, *Proc. Natl. Acad. Sci.* 109 (2012) 14381–14386.
- [10] J.A.L. Brown, J.K. Eykelboom, N.F. Lowndes, Co-mutation of histone H2AX S139A with Y142A rescues Y142A-induced ionizing radiation sensitivity, *FEBS Open Biol.* 2 (2012) 313–317.
- [11] K. Luger, T.J. Rechsteiner, A.J. Flaus, M.M.Y. Wayne, et al., Characterization of nucleosome core particles containing histone proteins made in bacteria, *J. Mol. Biol.* 272 (1997) 301–311.
- [12] L. Sie, S. Loong, E.K. Tan, Utility of lymphoblastoid cell lines, *J. Neurosci. Res.* 87 (2009) 1953–1959.
- [13] T. Maile, A. Izrael-Tomasevic, T. Cheung, G. Guler, et al., Mass spectrometric quantification of histone posttranslational modifications by a hybrid chemical labeling method, *Mol. Cell. Proteomics* 14 (2015) 1148–1158.
- [14] L. Matheron, H. van den Toorn, A.J.R. Heck, S. Mohammed, Characterization of biases in phosphopeptide enrichment by Ti4+-immobilized metal affinity chromatography and TiO₂ using a massive synthetic library and human cell digests, *Anal. Chem.* 86 (2014) 8312–8320.
- [15] S.D. Sherrod, M.V. Myers, M. Li, J.S. Myers, et al., Label-free quantitation of protein modifications by pseudo selected reaction monitoring with internal reference peptides, *J. Proteome Res.* 11 (2012) 3467–3479.
- [16] B. MacLean, D.M. Tomazela, N. Shulman, M. Chambers, et al., Skyline: an open source document editor for creating and analyzing targeted proteomics experiments, *Bioinformatics* 26 (2010) 966–968.
- [17] Y. Zheng, J.D. Tipton, P.M. Thomas, N.L. Kelleher, et al., Site-specific human histone H3 methylation stability: fast K4me3 turnover, *Proteomics* 14 (2014) 2190–2199.
- [18] J.V. Olsen, M. Vermeulen, A. Santamaria, C. Kumar, et al., Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis, *Sci. Signal.* 3 (2010) ra3.
- [19] K.T.G. Rigbolt, T.A. Prokhorova, V. Akimov, J. Henningsen, et al., System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation, *Sci. Signal.* 4 (2011) rs3.
- [20] A. Xie, A. Hartlerode, M. Stucki, S. Odate, et al., Distinct roles of chromatin-associated proteins MDC1 and 53BP1 in mammalian double-strand break repair, *Mol. Cell* 28 (2007) 1045–1057.